Biological Psychology 205 (2026) 109240

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Neural gender stereotype asymmetry in bidirectional word-face priming

Luana Serafini [a,1,*], Francesca Pesciarelli [a,b,2,**]

[a] Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Via Campi 287, Modena 41100, Italy
[b] Department of Psychology, San Diego State University (SDSU), CA, USA

ARTICLE INFO

Keywords:
Face processing
Gender stereotype
Event-related potentials (ERPs)
N400
P300/LPP

ABSTRACT

This study investigates the behavioral and neural correlates of implicit and automatic gender stereotyping at the intersection of language and face processing, focusing on the bidirectional flow of gender-stereotypical information and previously observed asymmetries in stereotype processing. We recorded response times (RTs) and Event-Related Potentials (ERPs) to a target face (male, female) preceded by stereotypically associated words (e.g., *conducente* "driver", *badante* "caregiver"), or to stereotypically associated target words preceded by a face (male, female). Participants performed a gender categorization task on target faces and a lexical decision task on target words/non-words. RTs showed typical priming effects for target faces, but an asymmetrical priming effect for target words: faster responses to stereotypically female words preceded by a gender-congruent face, whereas faster responses to stereotypically male words preceded by a gender-incongruent face. ERPs showed a gender stereotype asymmetry for target faces and only partially for target words. Female faces elicited a larger P300 and LPP when preceded by stereotypically gender-incongruent than -congruent prime words; male faces elicited a larger N400 when preceded by stereotypically gender-incongruent than -congruent prime words. Similarly, stereotypically male words elicited a larger P300 when preceded by a gender-incongruent than congruent face. Our findings reveal robust neural asymmetries in gender stereotype processing, extending beyond language, underscoring the need to treat female- and male-related biases as distinct cognitive phenomena in both future research and intervention.

Are women in male roles perceived the same as men in female roles? Growing neural evidence suggests asymmetrical processing of gender stereotypes. Studies employing event-related potentials (ERPs) to investigate the electrophysiological correlates of language-based gender stereotyping, have revealed distinct neural responses to violations of male and female stereotypes. The violation of female-related stereotypes by male agents elicited an N400 incongruence effect (Proverbio et al., 2018 in men; Siyanova-Chanturia et al., 2012; but see Weinstein et al., 2025 for a congruency effect), whereas the violation of male-related stereotypes by female agents elicited a P600 effect (Irmen et al., 2010; Proverbio et al., 2018 in men; Su et al., 2016). Studies that reported no asymmetry may have conflated male- and female-related stereotypes, highlighting the need to examine them as distinct phenomena. This pattern suggests that male- and female-related stereotypes activate through different mechanisms, at different timings. The N400 effect — a centro-parietal negative-going deflection occurring around 400 ms—could index the difficulty of accessing or retrieving the meaning of single words, or of integrating the meaning of a word in sentence contexts (Hagoort, 2007; Kutas et al., 2006). The P600 could index integration difficulties requiring a reanalysis and repair (Niharika & Prema Rao, 2020; Osterhout & Holcomb, 1992; 1995). Stereotype asymmetries may arise from the unmarked, generic use of the male gender (e.g., *uomo* "man," referring to both men and women), or from the unequal pace at which men and women have entered—or sought to enter—gender-atypical roles. While women are increasingly taking on roles traditionally held by men (Diekman & Eagly, 2000; World Bank, 2023), men rarely access roles traditionally held by women (England, 2010; Goldin, 2006). Social role theory (Eagly, 1987; Eagly & Steffen, 1984; Eagly & Wood, 2012) posits that gender stereotypes develop through an inferential process in which observed differences in role occupancy are

* Corresponding author.
** Corresponding author at: Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Via Campi 287, Modena 41100, Italy.
E-mail addresses: luana.serafini@unimore.it (L. Serafini), francesca.pesciarelli@unimore.it (F. Pesciarelli).
[1] https://orcid.org/0000-0002-7039-5154
[2] https://orcid.org/0000-0003-3007-4775

https://doi.org/10.1016/j.biopsycho.2026.109240
Received 10 October 2025; Received in revised form 4 March 2026; Accepted 6 March 2026
Available online 9 March 2026
0301-0511/

interpreted as reflecting stable dispositional traits. Changes in role distributions can foster stereotype change; however, given the unequal movement of women and men across roles, such change is likely to be asymmetric. This asymmetry could explain the different activation pattern observed here, whereby female-related stereotypes remain more rigid and their violation more severe, whereas male-related stereotypes show greater flexibility, with violations engaging only later adjustment processes. The Stereotype Content Model (SCM; Fiske et al., 2002; Fiske, 2018) further provides insight into additional contributing factors, such as status and attitudes. The model proposes that social groups are perceived along two fundamental dimensions—warmth and competence—which are systematically related to perceived social status. Women are more strongly associated with warmth-related traits, whereas men are more strongly associated with competence-related traits (e.g., Connor & Fiske, 2018; Ellemers, 2018). Competence is closely linked to higher perceived status and admiration, while warmth without competence is associated with lower status and paternalistic responses. Based on this model, we propose that men associated with female-related stereotypes could be less expected and elicit more negative evaluations, given the implied loss of perceived status; whereas women associated with male-related stereotypes could be more expected but still incur negative evaluations for violating prescriptive gender norms (Heyder & Kortzak, 2024; Tremmel & Wahl, 2023).

Our recent study (Serafini & Pesciarelli, 2025) demonstrated this neural asymmetry in face processing, thereby extending prior findings beyond language. Faces are far more salient and socially relevant than words, they convey gender cues rapidly (˜200 ms e.g., Ito & Urland, 2003; Mouchetant-Rostaing et al., 2000) and can automatically activate stereotypes (Ito & Tomelleri, 2017; Volpert-Esmond et al., 2017). In addition, pictures access semantic meaning more directly than words, and social category activation is thought to be stronger through person perception (Guilbeault et al., 2024; Macrae & Bodenhausen, 2000). These properties make them ecological and suitable for the investigation of automatic and implicit stereotyping. In our study, we recorded ERPs as Italian-speaking participants categorized by gender feminine and masculine Italian third-person singular pronouns (i.e., *lui* "he", *lei* "she"), and a female and a male face exemplar, primed by stereotypically associated common gender words (e.g. *insegnante* "teacher", *conducente* "driver") or grammatically marked female or male words with no stereotypical association (i.e. words ending by the feminine -a or masculine -o, e.g. *inesperto*$_{\text{MASC}}$ "inexpert", *antipatica*$_{\text{FEM}}$ "unpleasant"). When faces were primed by grammatically marked words, we observed both N400 and P300 effects. When faces were primed by gender stereotypical words, we found a gender asymmetry, consistent with the previously reported one: the priming effect was on the N400 for violations of the female stereotype by male faces (e.g., *badante* "caregiver" -male face), whereas on the P300 for violations of the male stereotype by female faces (e.g., *falegname* "carpenter"- female face). We did not find the typical ERP priming effects for pronouns, neither for the grammatical nor for the stereotypical condition, suggesting stronger gender effects for faces than words (for a detailed explanation see Serafini & Pesciarelli, 2025). Of note, in priming paradigms, P300-like components are often reported in place of the canonical P600 (Barber & Carreiras, 2003; Bentin et al., 1985; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). Compared to the classical P600, these effects are shorter-lived and show a well-defined peak, while also exhibiting a longer latency than the typical P300, possibly due to their temporal overlap with the preceding N400 (see Siyanova-Chanturia et al., 2012). In this context, the P300 is typically interpreted as indexing processes of reanalysis or context updating (Donchin, 1981).

The present study aimed to examine the reverse priming direction, i.e., stereotypically associated words primed by gender-conveying faces, as well as to replicate the previous asymmetry. Traditional models of person perception (e.g., Bodenhausen & Macrae, 1998; Brewer, 1988; Fiske & Neuberg, 1990) view stereotype activation as a bottom-up process, from member perception to categorization to stereotype activation: perceiving a woman would automatically activate the corresponding gender category and subsequently the female stereotype, so that stereotypic traits (e.g., "housewife," "caregiver") can become highly accessible. In everyday contexts, however, stereotypical information can flow in either direction: gender may be inferred from attributes, or attributes may be integrated once gender is known. The dynamic interactive theory of person construal (Freeman & Ambady, 2011) and other connectionist frameworks (e.g., Cox & Devine, 2015; McClelland et al., 2010) place category and attributes on equal footing, allowing for stereotype activation to also affect category activation (e.g., gender), so that our previous findings could be ascribed to the top-down influence of activated gender stereotypes on perception. Although complementary, these two conditions likely engage partially distinct processes, as the former requires category inference, whereas the latter involves the integration of stereotypical information with an already identified social category. One root may also be preferred for certain stereotypes, depending on which ordering is more consistently reinforced through experience (Cox & Devine, 2015). Presenting a face followed by a stereotypically associated word also allows us to test this more traditional pathway by examining how the brain responds to violations in the attribution of gender-stereotyped information to members of stereotyped groups.

In ERP language research, the order in which gender-related and stereotypical information is presented appears to influence sentence processing but not priming effects, as reported in the review by Porkert et al. (2024). In sentence comprehension studies, the N400 effect typically emerges when counter-stereotypical information must be integrated within a pre-activated gender category (e.g., "Li's son is a *nurse*") (Kreiner et al., 2009; Molinaro et al., 2016; Weinstein et al., 2025; Du & Zhang, 2023), whereas the P600 effect when stereotypical information inform gender inferences, such as during anaphoric resolution (e.g., "the doctor prepared *herself* for the operation") (Canal et al., 2015; Kreiner et al., 2009; Irmen et al., 2010; Osterhout et al., 1997; Su et al., 2016). Exceptions to this pattern are the findings by Weinstein et al. (2025) who contrarily reported a Late Positive Potential (LPP) effect when stereotypically associated words followed gender-conveying words, and Proverbio et al., (2017; 2018), in which the critical word was not fixed, but could be either the gender-revealing or the stereotype-conveying word. Gender stereotype asymmetries have been reported for both orders on the N400 (Grant et al., 2020; Proverbio et al., 2018; Weinstein et al., 2025) and P600 (Irmen et al., 2010; Su et al., 2016, Proverbio et al., 2018). Notably, to our knowledge, only one sentence processing study (Kreiner et al., 2009) directly compared both presentation orders within the same experiment, consistently finding a P600 when gender had to be inferred from the stereotypically associated word (e.g., "The *minister* left London after reminding *herself* about the letter"), and an early negativity in the N400 time window when it did not ("After reminding *herself* about the letter, the *minister* left London").

In contrast, priming paradigms with word pairs tend to elicit N400 effects for gender stereotype violations regardless of the direction of the priming, whether gender-conveying words follow stereotypically associated words (e.g., Driver – She) (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), or stereotypically associated words followed gender-conveying words (e.g., Men—Nurturing) (Wang et al., 2016; 2017; White et al., 2009). To date, asymmetries have been reported only in the former case (Siyanova-Chanturia et al., 2012).

Critically, no ERP study to date has directly compared both presentation orders within the same priming experiment in the context of gender stereotype processing. In our recent study combining words and faces within a priming paradigm (Serafini & Pesciarelli, 2025), we observed more complex ERP effects, including a P300 effect, suggesting that priming effects may be more nuanced than previously assumed. The greater complexity could be due to the use of faces or to the separate analysis of female and male targets. Indeed, faces have been shown to elicit both N400- and P300-like effects when preceded by stereotypical verbal or visual information (Rodríguez-Gómez et al., 2020; Sato et al., 2020; Serafini & Pesciarelli, 2025) and an N400-like effect when they prime stereotypical targets (Zhang et al., 2018), although asymmetries have only been reported when faces follow stereotype-conveying stimuli (Rodríguez-Gómez et al., 2020; Serafini & Pesciarelli, 2025).

The present study presents a novel condition and a replicated condition. In the replicated condition, face targets (male, female) were preceded by stereotypical words (e.g., *conducente* "driver", *badante* "caregiver"), exactly as in Serafini and Pesciarelli (2025). In the novel condition, the order was reversed: stereotypical word targets were preceded by faces. Italian speaking participants performed a gender categorization task on target faces, and a lexical decision task on target words/non-words. This design allows us to directly compare ERP effects across both presentation orders and examine whether the observed asymmetries and late positive effects are maintained when the sequence of information is inverted. These tasks focus on implicit processing without explicitly referencing gender-stereotype associations. We expected to replicate the face-related asymmetry: an N400 effect for violations of female stereotypes by male faces, and an P300 effect for violations of male stereotypes by female faces (Serafini & Pesciarelli, 2025). For word targets, we predicted the inverse pattern: an N400 effect when expectations of male-related information were violated by female-stereotypical words, and a P300 effect when expectations of female-related information were violated by male-stereotypical words. This prediction is based on the reported inverted stereotypical asymmetries found when stereotypical information needed to be integrated into a preactivated gender category (Weinstein et al., 2025). The effects might be stronger for words due to a more ecological priming order (see classical models of person perception cited above), or otherwise stronger for faces due to greater use of stereotypical information when gender can be inferred, or to the more implicit nature of the lexical decision task (White et al., 2018). In addition to the N400 and P300, we analyzed the LPP for target faces, following Serafini and Pesciarelli (2025) (see also Rodríguez-Gómez et al., 2020). The LPP is sensitive to contextual modulation in face processing and typically emerges between 400 and 700 ms over parietal–occipital electrodes (Cuthbert et al., 2000; Weinberg & Hajcak, 2010; Breton et al., 2014; Thom et al., 2014), reflecting enhanced attentional allocation to salient or contextually incongruent stimuli (Hajcak et al., 2010). Given these properties, we hypothesized that the LPP would be larger for faces primed by incongruent than congruent words.

1. Methods

1.1. Participants

Thirty-eight native Italian speakers participated in the experiment (20 women,[3] age range = 18–43 yrs, *M* = 21.29 yrs, *SD* = 4.37 yrs), none of which participated in our previous study (Serafini & Pesciarelli, 2025). Participants were undergraduate students from the University of Modena and Reggio Emilia, Italy, and researchers' acquaintances. Five female participants were excluded from all analyses due to low accuracy (< 90% overall in the face or word condition). The final sample consisted of 33 participants (15 women, two ambidextrous, age range = 18–43 yrs, *M* = 21.67 yrs, *SD* = 4.57 yrs). All participants had normal or corrected-to-normal vision, but one color blind, and they declared no history of neurological disorders. The final sample coincided with the sample size of our previous study (Serafini & Pesciarelli, 2025) and is in line with previous literature (Siyanova-Chanturia et al., 2012; Pesciarelli et al., 2019).

The experiment was approved by the local Ethical Committee (Comitato Etico dell'Area Vasta Emilia Nord- Italy; Protocol No. AOU 0008846/20) and it was run in accordance with the "Italian Association of Psychology" (AIP) Ethical Guidelines (Codice Etico: www.aipass.org/node/11560) and the Declaration of Helsinki. Participants provided written informed consent prior to participation, and university students received course credits.

1.2. Stimuli

Word stimuli were the same used in our previous works (Siyanova-Chanturia et al., 2012; Pesciarelli et al., 2019; Serafini & Pesciarelli, 2025). They consisted of 90 Italian common gender words conveying occupation, individual characteristics, or roles. Thirty were stereotypically female and 30 stereotypically male words (e.g., *insegnante* "teacher", *conducente* "driver", respectively); 30 were stereotypically neutral words, serving as fillers (e.g., *conoscente* "acquaintance"). Stereotypically female and male words received equally high evaluations of gender-oriented stereotypicality (see details in Siyanova-Chanturia et al., 2012). Fillers served to prevent participants from noticing the presence of the gender stereotypical words. All words ended either with -e or consonant, to avoid a gender-to-ending consistency typical of the Italian language (-a ending being typically feminine, -o ending being typically masculine). Ninety non-word stimuli were generated from gender-marked words previously rated as stereotypically neutral (see Siyanova-Chanturia et al., 2012 for details). For words shorter than seven letters, one letter was replaced; for words longer than eight letters, two letters were replaced (e.g., rip**aso**te, from riposato "rested"). All non-words were constrained to end in -e to reflect the most common word ending and to prevent participants from relying on sub-lexical strategies during the lexical decision task. Words were previously matched for written frequency, length and valence (see details in Siyanova-Chanturia et al., 2012). Non-words were matched to words in length.

Face stimuli were the same used in our previous work (Serafini & Pesciarelli, 2025) and consisted of one female and one male face exemplar, selected from the Chicago Face Database (Ma et al., 2015). The two exemplars resulted highly prototypical of the female and male gender categories (see details in Serafini & Pesciarelli, 2025) (see Fig. 1).

When faces appeared in the target position, each stereotypical prime word was paired with either a female or male face, resulting in two stereotypically congruent conditions (*pugile* "boxer" – male face, *colf* "maid" – female face) and two stereotypically incongruent conditions (*pugile* "boxer" – female face*, colf* "maid" – male face). When words appeared in the target position, each prime face was paired with a stereotypically male or female word, resulting in two stereotypically congruent conditions (male face – *pugile* "boxer", female face – *colf* "maid") and two stereotypically incongruent conditions (male face – *colf* "maid," female face – *pugile* "boxer"). Thus, each stereotypical word appeared four times throughout the whole experiment, twice in the prime position and twice in the target position. Each condition consisted of 30 trials, for a total of 240 experimental trials.

Each filler word was also paired with both the male and female face, and appeared both in the target position (*ospite* "guest" – male face, *ospite* "guest" – female face), and in the prime position (male face – *ospite* "guest", female face – *ospite* "guest"). Each filler condition consisted of 30 trials, for a total of 120 filler trials. Additionally, the male and female prime faces were paired with non-words, creating two non-word conditions (male face – *matore*, female face – *matore*). Each non-word condition consisted of 90 trials, for a total of 180 trials. In total, participants completed 540 trials.

1.3. Design and procedure

Participants were comfortably seated in a darkened, electrically shielded and sound-attenuated room. An example of the experimental procedure is illustrated in Fig. 1.

The procedure was adapted from Serafini and Pesciarelli (2025). In the center of the monitor, a fixation cross (+) appeared until the participant's response (note, in Serafini & Pesciarelli, it appeared for

[3] Participants' gender was assessed as gender identity, not biological sex.

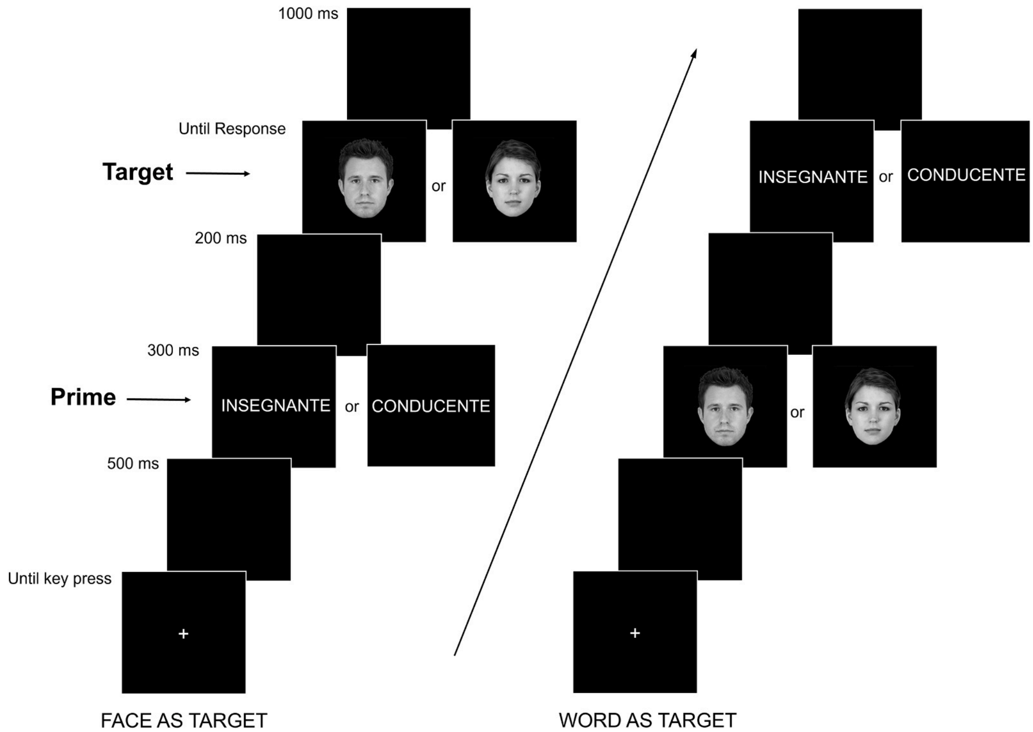

Fig. 1. Examples of the experimental procedure. Displayed are two conditions, one in which the face is the target (sequence on the left), and the other in which the word is the target (sequence on the right). In the first condition, the prime is a stereotypically associated word and the target is a face, which could be either gender-congruent (*insegnante* "teacher"- female face; *conducente* "driver"– male face) or gender-incongruent (*insegnante* "teacher"- male face; *conducente* "driver"- female face). In the second condition, the prime is a face and the target is a stereotypically associated word, which could be either gender-congruent (female face – *insegnante* "teacher"; male face – *conducente* "driver") or gender-incongruent (male face – *insegnante* "teacher"; female face – *conducente* "driver"). Displayed are the two face exemplars used in the experiment, together with two sample words from the stimulus set.

500 ms). Next, a blank screen appeared for 500 ms, then the prime word or face was displayed on the screen for 300 ms, followed by a blank screen for 200 ms. Subsequently, the target word (stereotypically male or female) or face (male or female) appeared and remained on the screen until the participant's response. A blank screen of 1000 ms followed each trial. To avoid a possible response automatism (see Serafini & Pesciarelli, 2025), participants were also explicitly instructed to attend to both primes and targets and to mentally read the prime words. Participants were instructed to decide, as quickly and as accurately as possible, whether the letter string was a word or a non-word (lexical decision task), or whether the face was male or female (gender categorization task). Participants responded by pressing one of two keys, counterbalanced (left, right) across participants, using their right and left indices.

Trials were organized in six blocks, three consecutive with target faces and three consecutive with target words/non-words, counterbalanced across participants. Blocks with target faces contained 60 trials each: 30 with female target faces, 30 with male target faces. Blocks with target words contained 120 trials each: 40 with stereotypical target words, 20 with filler target words, and 60 with target non-words. Repetitions of the same word appeared in different blocks and the blocks' order was counterbalanced across participants. Prime-target pairs were randomized within each block prior to presentation.

Before the experiment, participants completed two training sessions, one consisting of 12 prime-target pairs with word/non-word targets, and one consisting of six prime-target pairs with face targets. All conditions appeared once, with different stimuli than those used in the experimental session.

After the experiment, to measure individual stereotypical gender attitudes participants completed the *Bem Sex Role Inventory* (BSRI) (Gaudreau, 1977) and the *Ambivalence Sexism Inventory* (ASI) (Glick & Fiske, 1996) (see "Supplementary material" for analyses of these questionnaires).

1.4. EEG recording and analysis

EEG was recorded using the same system, setup and settings as in Serafini and Pesciarelli (2025) (see full details in the "Supplementary Material" of Serafini & Pesciarelli, 2025). In sum, EEG signal was recorded at a sampling rate of 1000 Hz with an ActiCHamp Plus (Brain Products GmbH, Gilching, Germany) system from 64 active electrodes (ActiCap Slim, Brain Products GmbH, Gilching, Germany). EEG signal was downsampled to 500 Hz, re-referenced to the average mastoid activity, and band-pass filtered from 0.01 to 80 Hz. Compromised channels were interpolated, and EEG signal was corrected for ocular and motion artifacts. A maximum of 3 channels were interpolated per participant (i.e., 4.69% of all channels). Continuous EEG signal was segmented into -200 to 800 ms epochs time-locked to target onset (instead of till 1000 ms, as in Serafini & Pesciarelli, 2025, given that effects were not observed beyond 800 ms and to retain more trials after artifact rejection) and 200-ms-baseline corrected. Epochs associated with correct responses were averaged across 8 conditions (mean number of epochs per condition = 28.27, SD = 0.56, range = 27.58–28.94). Lost data due to artifacts or incorrect responses was equal to 5.78% (SD = 2.76%).

For target face onset, based on visual inspection of aggregated grand-averaged ERP waveforms and in line with previous literature (e.g., Pesciarelli et al., 2019; Serafini & Pesciarelli, 2025; Siyanova-Chanturia et al., 2012), the following components were identified at frontal (F3, Fz, F4), central (C3, Cz, C4) and parietal (P3, Pz, P4) scalp sites: N400 from 250 to 400 ms after target onset; P300 from 300 to 500 ms after target onset; LPP from 600 to 800 ms after target onset. For target word onset, based on visual inspection of aggregated grand-averaged ERP waveforms and in line with previous literature (e.g., Pesciarelli et al., 2019; Serafini & Pesciarelli, 2025), the following components were identified at the same scalp sites: N400 from 250 to 450 ms after target onset; P300 from 450 to 650 ms after target onset. For each ERP component, amplitude was measured as mean activity within the respective time window.

1.5. Statistical analyses

We adopted a data-driven approach to leverage the rich spatiotemporal information of the EEG signal, providing a comprehensive perspective on the observed effects. To keep consistency with our previous study and the literature, we complemented the data-driven analyses with conventional ERP analyses performed on *a priori* defined time windows and electrodes (see Groppe et al., 2011).

Statistical analyses were performed in JASP (Version 0.17.2.1; JASP Team, 2022). Behavioral analyses were conducted only on trials with response latencies within ± 2 SD from each mean (rejected trials were on average 0.04%). Considering that face targets and word targets result in very different response and ERP latencies, we did not perform an omnibus ANOVA including the factor Target Type (face, word) but performed separate ANOVAs for face and word targets.

The mean RTs of correct responses per condition and the mean ERP amplitudes of correct trials were submitted to repeated-measure ANOVAs with the within-participant factors Congruency (congruent, incongruent) and Target Gender (male, female). Longitude (anterior, central, posterior) and Latitude (left, midline, right) were additional within-participant factors only for ERP analyses. The levels corresponded to the mean activity of F3, Fz, F4 (Anterior), C3, Cz, C4 (Central), P3, Pz, P4 (Posterior), F3, C3, P3 (Left), Fz, Cz, Pz (Midline) and F4, C4, P4 (Right). When Target Gender significantly interacted with Congruency, indicating a different priming effect depending on target gender (as predicted from our previous findings), separate ANOVAs for female and male targets were performed with the within-participant factor Prime Gender (male, female) and, for the EEG side, Latitude and Longitude.

We also conducted the same ANOVAs, including the between-participant factor Participant Gender (women, men), and, exclusively for the behavioral analyses, Task Order (target face first, target word first). Degrees of freedom were adjusted according to the Greenhouse-Geisser method; only corrected significance levels are reported. The level of significance testing was $p = .05$. Main effects and interactions that do not involve priming effects (i.e., not involving the effect of Congruency or Prime Gender for the ANOVAs separated for target gender) are not central to the research question and are therefore not reported. Hit rates were not analyzed due to ceiling effect, with all conditions averaging 95–99% correct.

We conducted mass univariate analyses in Brainstorm (Tadel et al., 2011), using cluster-based permutation tests to correct for multiple comparisons (Maris & Oostenveld, 2007), via the implemented FieldTrip function (Oostenveld et al., 2011). Individual averages per electrode per condition were imported in Brainstorm. Two-tailed paired *t*-tests were applied to compare incongruent and congruent conditions for female face target, male face target, stereotypically female word target, and stereotypically male word targets, at all electrode sites within the 200–800 ms post-stimulus time window. Clusters were formed by including adjacent electrodes and time points with *t*-scores exceeding ±2.36, corresponding to an uncorrected p-value of 0.05 for a single test, without requiring a minimum number of neighboring electrodes below this threshold to form a cluster. The spatial neighborhood for each electrode consisted of 5.4 neighboring electrodes per site on average. *T*-scores within each cluster were summed to compute a cluster-level *t*-score. Each cluster was assigned a *p*-value based on its ranking within the null distribution of the most extreme cluster-level *t*-scores, generated through 1000 permutations using a Monte Carlo approximation.

2. Results

2.1. Behavioral results

The repeated-measure ANOVA for target faces revealed a statistically significant main effect of congruency ($F(1,32) = 15.41$, $p < .001$, $\eta_p^2 = .325$, 90%CI [.112,.491]), indicating faster RTs when both the male and female face were preceded by congruent than incongruent stereotypical primes (see Table 1, Fig. 2).

The repeated-measure ANOVA for target words revealed a statistically significant Congruency x Target Gender interaction ($F(1,32) = 17.46$, $p < .001$, $\eta_p^2 = .353$, [.134,.515]). Separate ANOVAs for stereotypically male and female target words resulted in a significant prime gender effect for stereotypically female words ($F(1,32) = 7.56$, $p = .010$, $\eta_p^2 = .191$ [.028,.370]), indicating faster RTs when they were preceded by a congruent than incongruent prime face, and for stereotypically male words ($F(1,32) = 13.82$, $p < .001$, $\eta_p^2 = .302$ [.087,.463]), indicating a reversed priming effect, i.e., faster RTs when they were preceded by an *incongruent* than congruent prime face (see Table 1, Fig. 2).

When included in the analyses, the factors Participants' Gender and Task Order did not result in any significant effects (*ps* > .10).

2.2. ERP results from cluster-based permutation analyses

For face targets, the cluster-based permutation test indicated that there was a significant difference between the ERP amplitude elicited by the male target face primed by stereotypically incongruent and congruent words (cluster statistic (maxsum) = -2063, $p = .043$, cluster size = 805). The negative cluster which resulted statistically significant in the observed data extended from approximately from 280 to 380 ms post target mainly including central and parieto-occipital electrode sites (see Fig. 3). This effect is compatible for latency and spatial distribution with an N400 effect.

Additionally, the cluster-based permutation test indicated that there was a significant difference between the ERP amplitude elicited by the female target face primed by stereotypically incongruent and congruent words (cluster statistic (maxsum) = 2690, $p = .036$, cluster size = 1002). The positive cluster which resulted statistically significant in the observed data extended from approximately 600 to 750 ms post target and included frontal-central electrode sites (see Fig. 3). This effect is compatible for latency and spatial distribution with an LPP effect.

For word targets, the cluster-based permutation test indicated only a significant difference between the ERP amplitude elicited by stereotypically male target words primed by an incongruent as compared with a congruent face (cluster statistic (maxsum) = 2672, $p = .033$, cluster size = 1034). The positive cluster which resulted statistically significant in the observed data extended from approximately 500 to 700 ms post target and included central-parietal electrode sites (see Fig. 3). This effect is compatible for latency and spatial distribution with a P300 effect.

Table 1
Descriptive Statistics for Mean Correct Response Times (ms) in the Gender Categorization Task (Target Faces) and Lexical Decision Task (Target Words), by Prime Gender Congruency (Congruent, Incongruent).

Target	Prime Gender	*M*	*SD*	*95% CI*
Face				
Male	Congruent	488.91	76.08	462.95, 514.87
	Incongruent	510.36	69.52	486.64, 534.08
Female	Congruent	499.30	76.24	473.29, 525.32
	Incongruent	508.61	74.04	483.35, 533.87
Word				
Stereotypically female	Congruent	673.91	91.08	642.83, 704.98
	Incongruent	660.91	87.89	630.92, 690.90
Stereotypically male	Congruent	646.06	79.44	618.96, 673.16
	Incongruent	660.94	96.16	628.13, 693.75

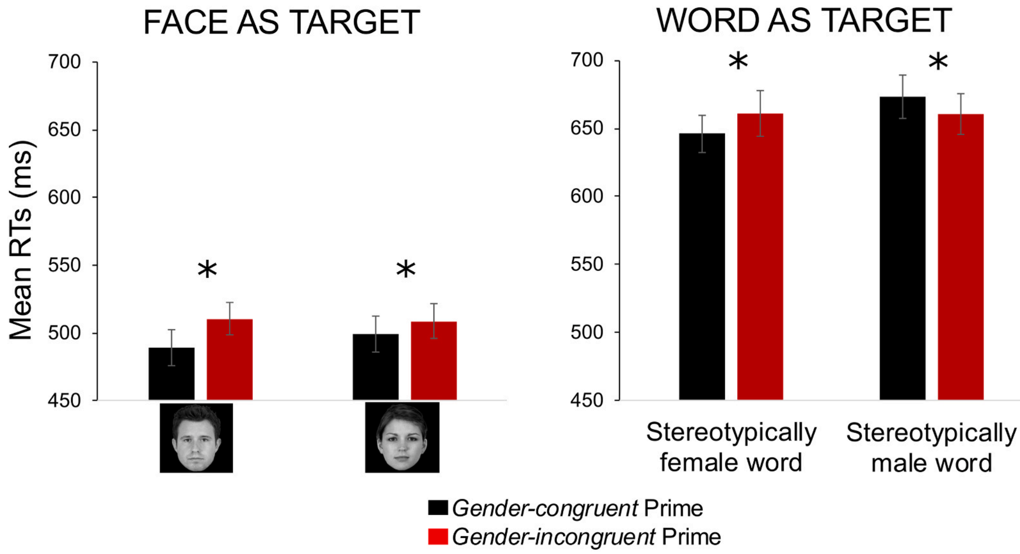

Fig. 2. Behavioral results. The bar plot on the left shows response times (RTs) to correctly categorize the male and female target face by gender, when preceded by stereotypically gender-congruent prime words (black/dark grey) or gender-incongruent prime words (red/light grey). The bar plot on the right shows RTs to correctly categorize stereotypically female and male target words as words, when preceded by a gender-congruent prime face (black/dark grey) or a gender-incongruent prime face (red/light grey). Asterisks (*) mark significant differences ($p < .05$).

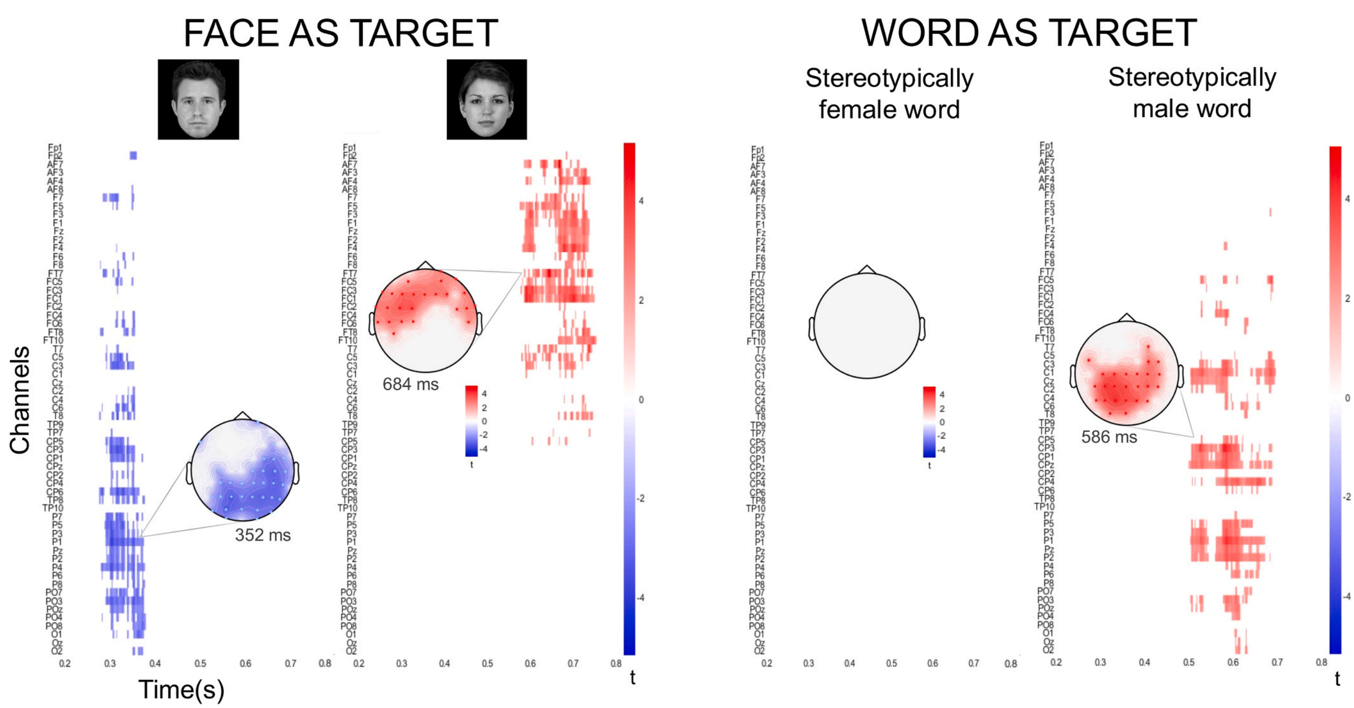

Fig. 3. ERP results from cluster-based permutation analyses. Raster plots show statistically significant differences between gender-incongruent and gender-congruent ERPs when a face was the target (left) and when a word was the target (right). Blue and red rectangles indicate electrodes/time points where incongruent ERPs were more negative or more positive than congruent ERPs, respectively; white indicates no statistically significant differences. Electrodes are ordered frontal to posterior, and the time window spans 200–800 ms post-target onset. Topographical maps depict incongruent-minus-congruent effects at the most representative time point for each condition.

2.3. ERP results from conventional analyses

2.3.1. Face targets

For the mean amplitude of the N400 (250–400 ms), the repeated-measure ANOVA revealed a statistically significant Congruency x Target Gender interaction ($F(1,32) = 5.75$, $p = .022$, $\eta_p^2 = .152$, 90% CI [.012,.330]). Separate ANOVAs for male and female target faces showed a significant prime gender effect ($F(1,32) = 5.35$, $p = .027$, $\eta_p^2 = .143$ [.009,.321]) only for male target faces, indicating larger N400 when they were primed by stereotypically incongruent than congruent words (see Table 2, Fig. 4).

For the mean amplitude of the P300 (300–500 ms), the repeated-measure ANOVA revealed a statistically significant Congruency x Target Gender x Longitude interaction ($F(2,64) = 6.25$, $p = .007$, $\varepsilon = .79$, $\eta_p^2 = .163$, [.029,.300]). Separate ANOVAs for male and female target faces showed a significant Prime Gender x Longitude interaction

Table 2
Descriptive Statistics for Mean ERP Amplitudes (μV) to Target Faces (Male, Female) by Prime Gender Congruency (Congruent, Incongruent) in the N400, P300, and LPP Time Windows across all Electrode Considered.

Target Gender	Prime Gender	*M*	*SD*	*95% CI*
N400				
Male	Congruent	4.26	1.96	3.57, 4.96
	Incongruent	3.79	2.05	3.06, 4.51
Female	Congruent	3.57	2.26	2.77, 4.38
	Incongruent	3.71	2.08	2.97, 4.44
P300				
Male	Congruent	1.95	2.10	1.20, 2.70
	Incongruent	1.73	2.22	0.94, 2.51
Female	Congruent	1.78	2.47	0.90, 2.65
	Incongruent	2.16	2.22	1.38, 2.95
LPP				
Male	Congruent	2.00	2.08	1.26, 2.73
	Incongruent	1.77	2.23	0.98, 2.56
Female	Congruent	1.76	2.47	0.88, 2.63
	Incongruent	2.15	2.23	1.36, 2.94

(F(2,64) = 5.87, p = .013, ε = .66, η_p^2 = .155 [.019,.307]) only for female target faces, indicating larger P300 when they were primed by stereotypically incongruent than congruent words, at anterior sites (t(32) = -2.82, p = .049) (see Table 2, Fig. 4).

For the mean amplitude of the LPP (600–800 ms), the repeated-measure ANOVA revealed a statistically significant Congruency x Target Gender x Longitude interaction (F(2,64) = 5.99, p = .008, ε = .78, η_p^2 = .158, [.025,.295]). Separate ANOVAs for male and female target faces showed a significant Prime Gender x Longitude interaction (F(2,64) = 7.18, p = .005, ε = .69, η_p^2 = .183 [.035,.333]) only for female target faces, indicating larger LPP when they were primed by stereotypically incongruent than congruent words, at anterior sites (t(32) = -2.94, p = .040) (see Table 2, Fig. 4).

When included in the analyses, the factor Participants' Gender did not result in any significant effects (ps > .10).

2.3.2. Word targets

For the mean amplitude of the N400 (250–450 ms), the repeated-measure ANOVA revealed a statistically significant Congruency x Target Gender interaction (F(1,32) = 5.43, p = .026, η_p^2 = .145, [.010,.323]), but separate ANOVAs for stereotypically female and male target words showed no significant prime gender effects (ps > .10). The analysis also revealed a significant Congruency x Latitude x Longitude interaction (F(4128) = 2.67, p = .043, ε = .87, η_p^2 = .077, [.001,.142]), but post-hoc comparisons did not reveal any significant priming effects (all ps >.10) (see Table 3, Fig. 4).

When the factor Participants' Gender was included in the analysis, a statistically significant Congruency x Longitude x Latitude x Participant's Gender interaction emerged (F(4124) = 3.23, p = .022, ε = .83, η_p^2 = .094 [.008,.168]), but post-hoc comparisons did not reveal any significant priming effects (all ps >.10).

For the mean amplitude of the P300 (450–650 ms), the repeated-measure ANOVA revealed a statistically significant Congruency x Target Gender interaction (F(1,32) = 4.84, p = .035, η_p^2 = .131, [.005,.308]). Separate ANOVAs for stereotypically male and female target words showed a significant Prime Gender effect (F(1,32) = 4.92, p = .034, η_p^2 = .133 [.006,.310]) only for stereotypically male words, indicating larger P300 when they were primed by a gender-incongruent than -congruent face (see Table 3, Fig. 4).

When included in the analysis, the factors Participants' Gender did not result in any significant effects (ps > .10).

3. Discussion

The aim of the present study was two-fold: (i) to investigate whether the neural gender stereotype asymmetry found in our previous word-face priming paradigm (Serafini & Pesciarelli, 2025) was maintained by inverting the order of the information stream, using a reversed face–word priming paradigm, and (ii) to replicate the asymmetry using the previous presentation order. To our knowledge, the present study represents the first ERP investigation combining both presentation orders within the same priming study in the context of gender stereotypes.

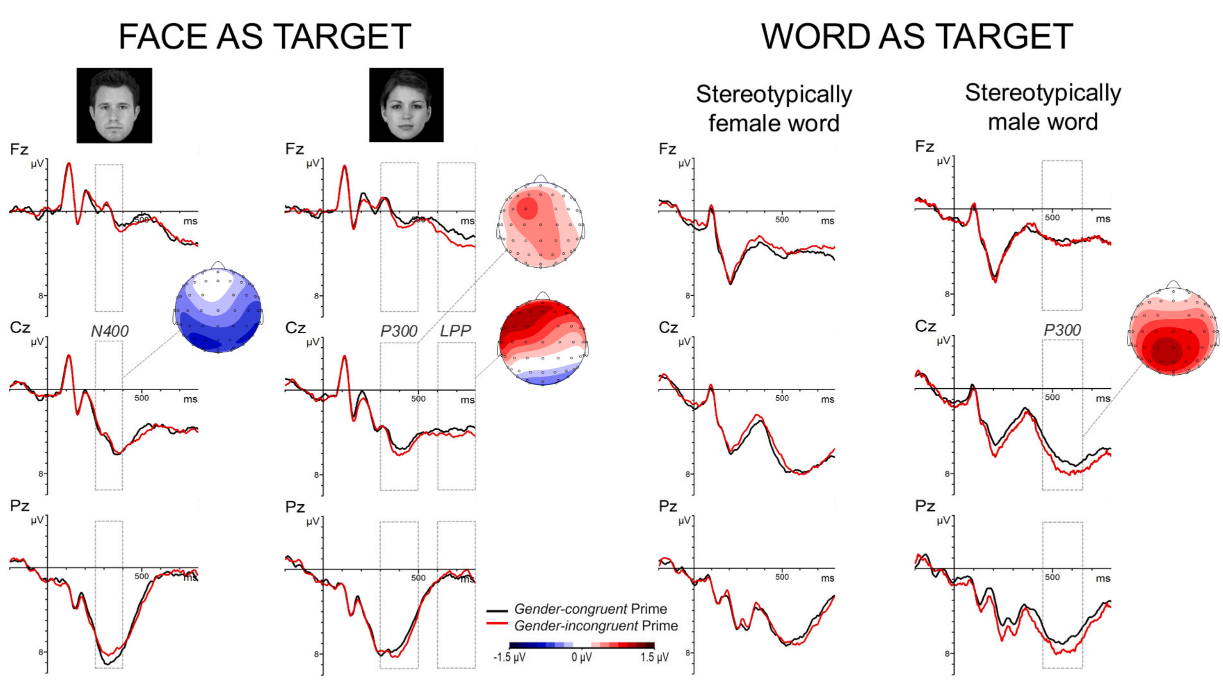

Fig. 4. ERP results from conventional analyses. Plots illustrate the grand-averaged ERP waveforms to the face target (left) and word target (right), primed by gender-congruent (black/dark grey) or gender-incongruent (red/light grey) primes. Waveforms are shown for electrodes Fz, Cz, and Pz, representative of the observed effects, and are time-locked to the target onset (negative voltages are plotted upward). Topographical maps illustrate the scalp distribution of the mean incongruent-minus-congruent effect for each time window and condition showing a statistically significant effect (p < .05).

Table 3
Descriptive Statistics for Mean ERP Amplitudes (μV) to Target Words (Stereotypically Female, Stereotypically Male) by Prime Gender Congruency (Congruent, Incongruent) in the N400 and P300 Time Windows across all Electrode Considered.

Target Gender	Prime Gender	*M*	*SE*	*95% CI*
N400				
Stereotypically female	Congruent	2.75	2.51	1.86, 3.65
	Incongruent	3.17	2.36	2.34, 4.01
Stereotypically male	Congruent	3.51	2.23	2.72, 4.30
	Incongruent	3.20	2.33	2.37, 4.03
P300				
Stereotypically female	Congruent	4.34	2.70	3.38, 5.29
	Incongruent	4.93	2.39	4.08, 5.78
Stereotypically male	Congruent	5.01	2.19	4.24, 5.79
	Incongruent	4.92	2.25	4.12, 5.72

We recorded RTs and ERPs while participants performed a lexical decision task on stereotypically female and male words and non-words, primed by female and male faces (novel condition), or categorized by gender female and male faces, primed by stereotypically female or male words (replicated condition). Results revealed a novel consistent but partial ERP gender stereotype asymmetry when words were targets, and the expected asymmetry when faces were targets.

Indeed, at both behavioral and neural levels, we successfully replicated our earlier findings concerning gender stereotype priming effects for face targets. Behaviorally, face gender categorization was faster when faces were preceded by gender-congruent primes, echoing established findings from both linguistic priming (e.g., Banaji & Hardin, 1996; Cacciari & Padovani, 2007; Casado et al., 2023; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), and facial non-priming research (e.g., Becker et al., 2007; Dzhelyova et al., 2012; Le Gal & Bruce, 2002; Fitousi, 2021; Zhu et al., 2020). The ERP effects also replicated a gender-based asymmetry. Both data-driven and conventional analyses showed that violations of female-related stereotypes by a male face (e.g., *badante* "caregiver"-male face) elicited an early N400 effect, consistent with the idea of a semantic expectancy violation (Kutas & Hillyard, 1980; Kutas & Federmeier, 2011), while violations of male-related stereotypes by a female face (e.g., *falegname* "carpenter"- female face) elicited a later P300/LPP effect, possibly indicating enhanced evaluative processing, or perceived improbability and context updating (Donchin, 1981; Duncan-Johnson & Donchin, 1977; Sutton & Ruchkin, 1984). These results again align with previous ERP findings in language-exclusive tasks (e.g.,Irmen et al., 2010; Proverbio et al., 2018 in men; Siyanova-Chanturia et al., 2012; Su et al., 2016; but see Weinstein et al., 2025 for a congruence effect), and with Rodríguez-Gómez et al. (2020)'s study using both words and faces, finding a seemingly more complex processing of female-related gender stereotype as indexed by various LPP effects. However, they do not align with Rodríguez-Gómez et al. (2020)'s findings of a larger LPP for male faces preceded by sentences conveying female stereotypes, a discrepancy that may reflect the more explicit nature of their task, which likely engaged stronger evaluative processes.

While the asymmetry itself replicates our previous results, there were differences in the temporal and topographical properties of the effects. The N400 here exhibited a more centro-parietal distribution, in line with classic semantic processing studies and our previous studies using target pronouns (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012), whereas in the replicated study we observed a more fronto-central distribution. The late positivity, while again present, showed a longer latency (600–750 ms), and a more frontal distribution as compared to the previously observed centro-parietally distributed one. The slightly slower response latencies compared to our previous study (Serafini & Pesciarelli, 2025) likely reflect the successful reduction of automatic response tendencies due to procedural changes (e.g., reintroducing a spacebar press to start each trial, emphasizing the importance of reading the prime words). They also aligned with those observed in previous priming studies on pronouns (Cacciari & Padovani, 2007; Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012). It is possible that the more typical response latencies contributed to these ERP patterns, which may reflect a more traditional processing sequence when response automatism is minimized. However, further replication is needed to confirm this interpretation. In this study as compared to the previous, the positive effect appears to reflect an LPP rather than a P300, which can be differently interpreted as enhanced effort and/or attentional engagement during evaluative processing (Hajcak & Foti, 2020), possibly indicative of an implicit negative evaluation of female faces presented in male-stereotypical contexts. However, the ANOVA also revealed a P300 effect alongside the LPP, possibly due to the increased sensitivity of the *a priori* analysis. Moreover, the literature suggests that the P300 and LPP may index the same underlying process, namely context updating (e.g., Fields, 2023). Overall, the replication reinforced the evidence for a gender stereotype asymmetry in face processing, in line with the social role theory (Eagly, 1987; Eagly & Wood, 2012) and the Stereotype Content Model (Fiske et al., 2002; Fiske, 2018) of person perception.

The face–word priming condition represents the novelty of this study. While prior priming studies have investigated the neural correlates of gender stereotyping using faces as targets, no ERP research to date has examined the reverse direction: the neural processing of stereotype-related words primed by congruent or incongruent facial gender cues. At the behavioral level, we found an unexpected gender stereotype asymmetry. Stereotypically female words were categorized faster as belonging to the word category when preceded by gender congruent face primes (e.g., female face– *badante* "caregiver" vs. male face– *badante* "caregiver). However, for stereotypically male words an inverse priming effect emerged: faster responses occurred following gender *incongruent* face primes (e.g., female face- *falegname* "carpenter" vs. male face- *falegname* "carpenter"). This reversal is atypical and not easily explained by conventional priming mechanisms (see Habibi & Khurana, 2012, exp.2; Zhang et al., 2019, exp.1, 3 for typical priming effects in similar studies). One interpretation of our findings is that participants may have consciously or unconsciously inferred the study's aims and responded in ways that reflect evolving societal attitudes—namely, greater familiarity or comfort with women in traditionally male roles. This awareness could be justified by the absence of grammatical control conditions, despite the inclusion of filler trials. This interpretation aligns with the one provided for the inverse ERP priming reported by Rodríguez-Gómez et al. (2020), who used a more explicit task. Despite reversed behavioral priming effects have not yet been reported in gender stereotype contexts, behavioral asymmetries have nonetheless emerged, showing more consistent priming when female stereotypes were violated by male referents (Chalyvidou & Weber, 2025; Reali et al., 2015), or an inhibition of the processing of male pronouns after female stereotypes (Cacciari & Padovani, 2007). The results from our study fits into this literature, overall suggesting heightened sensitivity to violations of female stereotypes. This finding is again consistent with social role theory (Eagly, 1987; Eagly & Wood, 2012) and can be interpreted as reflecting a more rigid female-related stereotype relative to a more flexible male-related stereotype.

An alternative interpretation concerns the use of a lexical decision task, which has yielded mixed results in the gender stereotype priming literature (Kidder et al., 2018; White et al., 2018). White et al. (2018) reported finding no effects at all in lexical decision paradigms, whereas Kidder et al.'s (2018) meta-analysis showed stronger effects on tasks that explicitly focused on gender. Together, these studies suggest that stereotype priming may depend on both semantic activation and response-related processes, similarly to evaluative priming (Herring et al., 2013). In gender categorization tasks, primes activate not only gender-related semantic information but also corresponding response tendencies, as the prime both pre-activates the gender category and allows for response preparation. In contrast, lexical decision tasks primarily engage semantic processing, as prime gender information is not relevant for preparing the required response (word vs. nonword). We believe that this variability warrants caution in interpreting the present effect and motivates future replication using a more gender-focused task.

Underlying this atypical behavioral effect, both data driven and conventional ERP analyses showed that stereotypically male words elicited a larger P300 when primed by a gender-incongruent than congruent face (e.g., female face - *falegname* "carpenter" vs. male face - *falegname* "carpenter"). No corresponding ERP effect was found in response to stereotypically female words, condition in which the behavioral priming showed the typical pattern. This effect aligns with Weinstein et al. (2025), who also reported an LPP for stereotypically associated words following descriptive gender cues, but diverges from most sentence processing, face processing, and priming studies, which typically report N400 effects for counter-stereotypical integration (e.g., Kreiner et al., 2009; Molinaro et al., 2016; Wang et al., 2016, 2017; Zhang et al., 2018). When faces are included and asymmetries are considered, gender-stereotype priming appears more nuanced than a purely N400-based account, involving an additional positive component not previously shown to be sensitive to stereotype violations.

Critically, our ERP findings on the target words revealed a partial gender stereotype asymmetry. This constitutes a novel finding, extending previous knowledge, by examining for the first time the ERP correlates of the processing of stereotype-conveying words primed by female and male faces. Gender stereotype asymmetries on late positive effects so far only emerged when the target was the gender-conveying word (Irmen et al., 2010; Rodríguez-Gómez et al., 2020; Serafini & Pesciarelli, 2025; Su et al., 2016, but see Proverbio et al., 2018 with mixed presentation order). When gender cues preceded stereotyped content, a gender stereotype asymmetry was only reported by Weinstein et al. (2025). They contrarily found a larger N400 effect for female-related rather than male-related adjectives following male referents (e.g., *hombre–sensible* "man- sensitive" vs. *hombre–fuerte,* "man – strong") –but no such effect for the female referent. As for Weinstein et al., who found the N400 effect (i.e., a *negativity*) in response to the *female*-related adjective, we also found an apparently reversed effect (i.e., more *positive* for *male* targets) as compared to the one we found on target faces. We argue that it could reflect a similar mechanism as for the reversed order: lower perceived probability of ascribing stereotypically male attributes to women than to men, as for the interpretation of a P300 (Donchin, 1981; Duncan-Johnson & Donchin, 1977; Sutton & Ruchkin, 1984). However, rather than indexing a reanalysis or context update—which would be expected to slow word processing—this effect may instead reflect an increased allocation of attention or processing resources to these words, resulting in faster responses when they were primed by incongruent rather than congruent faces. This heightened sensitivity to stereotypically male words following a female face may stem from participants' awareness of the study's scope and is consistent with social role theory (Eagly, 1987; Eagly & Wood, 2012).

Notably, our ERP study investigated both presentation orders within the same priming experiment, which was so far not attempted. We showed that, as the separate literature suggested, the priming paradigm worked similarly irrespective of the order of the information stream (Pesciarelli et al., 2019; Siyanova-Chanturia et al., 2012; Wang et al., 2016; 2017; White et al., 2009), which appear not to be the case for sentence comprehension paradigms that are more sensitive to the information order (see Porkert et al., 2024 for a review). The similarity of effects across presentation orders supports recent connectionist frameworks that assume bidirectional activation of stereotypical knowledge (Cox & Devine, 2015; McClelland et al., 2010; Freeman & Ambady, 2011). However, meaningful asymmetries also emerged across presentation orders: canonical priming effects were observed when words served as primes and faces as targets, whereas effects were atypical or partial when faces primed words. This asymmetry may reflect stronger stereotyping in one direction (as also predicted by connectionist frameworks), or depend on task demands: gender-stereotypical information is more readily used to infer a target's gender in categorization tasks, but facial gender cues may be downweighted in lexical decision tasks when they are not task-relevant. While the word–face and face–-word conditions are not fully comparable, given that they differ in stimulus variability (e.g., one face vs. multiple words per category) and task implicitness, it is nonetheless notable that they revealed a comparable asymmetry in gender stereotype activation. Critically, observing this effect even under less explicitly gender-focused tasks (i.e., a lexical decision task) reinforces the view that such biases operate automatically, beyond conscious intention or task demands, a claim further supported by our masked priming findings (Pesciarelli et al., 2019).

A promising direction for future research is to examine how individual differences influence ERP asymmetries, particularly N400 versus P300/LPP dominance, or P300 versus LPP dominance, following approaches such as Beatty-Martínez et al. (2021). In exploratory analyses (see Supplementary material), we assessed the role of self-ascribed masculinity and femininity (BSRI) and reported sexism (ASI). Overall, BSRI scores influenced ERP priming effects in the P300/LPP time window for target faces—especially female faces—and in the N400 time window for target words. These findings suggest that P300/LPP effects for faces may be sensitive to individual differences, whereas N400 effects for stereotypical words could emerge in individuals with higher self-ascribed masculinity traits.

To conclude, this study provides compelling evidence for neural gender stereotype asymmetries, which consistently emerge beyond language processing, across replications, and regardless of the order of information presentation—strongly suggesting that female- and male-related stereotypes' activation engages different cognitive mechanisms. This evidence has significant implications for how we study, understand, and ultimately intervene in gender bias, highlighting the need to consider female- and male-related stereotypes as distinct phenomena.

CRediT authorship contribution statement

Francesca Pesciarelli: Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Luana Serafini:** Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors did not use generative AI technologies for preparation of this work. Generative AI was not used during the preparation of this manuscript.

Study funding

Funded by the PNRR - Missione 4 'Istruzione e Ricerca' - Componente C2 Investimento 1.1 'Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN)', 'Gender bias in language: Testing INClusive ITalian language feasibility and impact (INCITƏ)', codice progetto 202272BY39 (CUP E53D23011740006). Funding sources had no involvement in the study design, collection, analysis and interpretation of data, writing of the report and decision to submit the article for publication.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to Andrea Di Rienzo, Alessia Salvato, Elena Zanichelli and Dr. Eleonora Borelli for assistance with participants' recruitment and data acquisition.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.biopsycho.2026.109240.

Data Availability

Processed data, aggregated per participant and condition, the detailed pre-processing pipeline description, and statistical analyses in JASP files are available at osf: https://osf.io/nfdms/?view_only=3da2e0dd453d418e84aa7e41afa41309.

References

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7*(3), 136–141. https://doi.org/10.1111/j.1467-9280.1996.tb00346.x

Barber, H., & Carreiras, M. (2003). Integrating gender and number information in Spanish word pairs: An ERP study. *Cortex; a Journal devoted to the Study of the Nervous System and Behavior, 39*(3), 465–482. https://doi.org/10.1016/s0010-9452(08)70259-4

Beatty-Martínez, A. L., Bruni, M. R., Bajo, M. T., & Dussias, P. E. (2021). Brain potentials reveal differential processing of masculine and feminine grammatical gender in native Spanish speakers. *Psychophysiology, 58*(3), Article e13737. https://doi.org/10.1111/psyp.13737

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology, 92*(2), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*(4), 343–355. https://doi.org/10.1016/0013-4694(85)90008-2

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr (Ed.), *Stereotype activation and inhibition* (pp. 1–52). Lawrence Erlbaum Associates Publishers.

Breton, A., Jerbi, K., Henaff, M. A., Cheylus, A., Baudouin, J. Y., Schmitz, C., Krolak-Salmon, P., & Van der Henst, J. B. (2014). Face the hierarchy: ERP and oscillatory brain responses in social rank processing. *PloS One, 9*(3), Article e91451. https://doi.org/10.1371/journal.pone.0091451

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull, & R. S. Wyer, Jr. (Eds.), *A dual process model of impression formation* (pp. 1–36). Lawrence Erlbaum Associates, Inc.

Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics, 28*(2), 277–293. https://doi.org/10.1017/S0142716407070142

Canal, P., Garnham, A., & Oakhill, J. (2015). Beyond Gender Stereotypes in Language Comprehension: Self Sex-Role Descriptions Affect the Brain's Potentials Associated with Agreement Processing. *Frontiers in Psychology, 6*, 1953. https://doi.org/10.3389/fpsyg.2015.01953

Casado, A., Sá-Leite, A. R., Pesciarelli, F., & Paolieri, D. (2023). Exploring the nature of the gender-congruency effect: implicit gender activation and social bias. *Frontiers in Psychology, 14*, Article 1160836. https://doi.org/10.3389/fpsyg.2023.1160836

Chalyvidou, D., & Weber, A. (2025). The role of grammatical gender and gender stereotypes in noun processing: The Tug of War in Greek. *Social Psychological Bulletin, 20*, 1–26. https://doi.org/10.32872/spb.13469

Connor, R. A., & Fiske, S. T. (2018). Warmth and competence: A feminist look at power and negotiation. In In. C. B. Travis, J. W. White, A. Rutherford, W. S. Williams, S. L. Cook, & K. F. Wyche (Eds.), *APA handbook of the psychology of women: History, theory, and battlegrounds* (pp. 321–342). American Psychological Association. https://doi.org/10.1037/0000059-016.

Cox, W. T., & Devine, P. G. (2015). Stereotypes possess heterogeneous directionality: a theoretical and empirical exploration of stereotype structure and content. *PloS One, 10*(3), Article e0122292. https://doi.org/10.1371/journal.pone.0122292ù

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology, 52*(2), 95–111. https://doi.org/10.1016/s0301-0511(99)00044-7

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin, 26*(10), 1171–1188. https://doi.org/10.1177/0146167200262001

Donchin, E. (1981). Surprise! … Surprise? *Psychophysiology, 18*(5), 493–513. https://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Du, Y., & Zhang, Y. (2023). Strategic processing of gender stereotypes in sentence comprehension: An ERP study. *Brain Sciences, 13*(4), 560. https://doi.org/10.3390/brainsci13040560

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology, 14*(5), 456–467. https://doi.org/10.1111/j.1469-8986.1977.tb01312.x

Dzhelyova, M., Perrett, D. I., & Jentzsch, I. (2012). Temporal dynamics of trustworthiness perception. *Brain Research, 1435*, 81–90. https://doi.org/10.1016/j.brainres.2011.11.043

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum Associates, Inc.

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage Publications Ltd. https://doi.org/10.4135/9781446249222.n49.

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology, 69*, 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

England, P. (2010). The gender revolution: Uneven and stalled. *Gender & Society, 24*(2), 149–166. https://doi.org/10.1177/0891243210361475

Fields, E. C. (2023). The P300, the LPP, context updating, and memory: What is the functional significance of the emotion-related late positive potential? *International Journal of Psychophysiology, 192*, 43–52.

Fiske, S. T. (2018). Stereotype content: warmth and competence endure. *Current directions in psychological Science, 27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of personality and Social Psychology, 82*(6), 878–902.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology, 23* pp. 1–74). Academic Press. https://doi.org/10.1016/S0065-2601(08)60317-2.

Fitousi, D. (2021). Stereotypical Processing of Emotional Faces: Perceptual and Decisional Components. *Frontiers in Psychology, 12*, Article 733432. https://doi.org/10.3389/fpsyg.2021.733432

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*(2), 247–279. https://doi.org/10.1037/a0022327

Gaudreau, P. (1977). Factor analysis of the Bem Sex-Role Inventory. *Journal of Consulting and Clinical Psychology, 45*(2), 299–302. https://doi.org/10.1037/0022-006X.45.2.299

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

Goldin, C. (2006). The Quiet Revolution that Transformed Women's Employment, Education, and Family. *American Economic Review, 96*, 1–21. https://doi.org/10.1257/000282806777212350

Grant, A., Grey, S., & van Hell, J. G. (2020). Male fashionistas and female football fans: Gender stereotypes affect neurophysiological correlates of semantic processing during speech comprehension (Article) *Journal of Neurolinguistics, 53*, Article 100876. https://doi.org/10.1016/j.jneuroling.2019.100876.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Guilbeault, D., Delecourt, S., Hull, T., Desikan, B. S., Chu, M., & Nadler, E. (2024). Online images amplify gender bias. *Nature, 626*(8001), 1049–1055. https://doi.org/10.1038/s41586-024-07068-x

Habibi, R., & Khurana, B. (2012). Spontaneous gender categorization in masking and priming studies: key for distinguishing Jane from John Doe but not Madonna from Sinatra. *PloS One, 7*(2), Article e32377. https://doi.org/10.1371/journal.pone.0032377

Hagoort, P. (2007). The memory, unification, and control (MUC) model of language. In In. A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing* (pp. 243–270). Psychology Press.

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology, 57*(7), Article e13570. https://doi.org/10.1111/psyp.13570

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental neuropsychology, 35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: a quarter century of evaluative priming research. *Psychological Bulletin, 139*(5), 1062–1089. https://doi.org/10.1037/a0031309

Heyder, A., & Kortzak, O. (2024). Backlash Against Women and Support for Men? On the Perception of Men and Women Researchers in Gender Stereotype-Inconsistent Domains. In *Psychology of Women Quarterly, 48* pp. 411–429). Original work published 2024. https://doi.org/10.1177/03616843241236397

Irmen, L., Holt, D. V., & Weisbrod, M. (2010). Effects of role typicality on processing person information in German: evidence from an ERP study. *Brain Research, 1353*, 133–144. https://doi.org/10.1016/j.brainres.2010.07.018

Ito, T. A., & Tomelleri, S. (2017). Seeing is not stereotyping: The functional independence of categorization and stereotype activation. *Social Cognitive and Affective Neuroscience, 12*(5), 758–764. https://doi.org/10.1093/scan/nsx009

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of*

Personality and Social Psychology, 85(4), 616–626. https://doi.org/10.1037/0022-3514.85.4.616
2022 JASP Team. (2022). JASP (Version 0.17.2.1).
Kidder, C. K., White, K. R., Hinojos, M. R., Sandoval, M., & Crites, S. L., Jr (2018). Sequential Stereotype Priming: A Meta-Analysis. *Personality and Social Psychology review: an Official Journal of the Society for Personality and Social Psychology, Inc, 22* (3), 199–227. https://doi.org/10.1177/1088868317723532
Kreiner, H., Mohr, S., Kessler, K., & Garrod, S. (2009). Can Context Affect Gender Processing? ERP Differences Between Definitional and Stereotypical Gender. In K. Alter, M. Horne, M. Lindgren, & J. von Koss Torkildsen (Eds.), *Brain Talk: Discourse With and in the Brain* (pp. 107–119). Lund: Media-Tryck.
Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science (New York, N. Y.), 207*(4427), 203–205. https://doi.org/10.1126/science.7350657
Kutas, M., Van Petten, C., & Kluender, M. (2006). Psycholinguistics electrified II. In M. A. Gernsbacher, & M. Traxler (Eds.), *Handbook of psycholinguistics* (2nd ed.). New York: Elsevier Press.
Le Gal, P. M., & Bruce, V. (2002). Evaluating the independence of sex and expression in judgments of faces. *Perception & psychophysics, 64*(2), 230–243. https://doi.org/10.3758/bf03195789
Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5
Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: thinking categorically about others. *Annual Review of Psychology, 51*, 93–120. https://doi.org/10.1146/annurev.psych.51.1.93
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024
McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences, 14*(8), 348–356. https://doi.org/10.1016/j.tics.2010.06.002
Molinaro, N., Su, J. J., & Carreiras, M. (2016). Stereotypes override grammar: Social knowledge in sentence comprehension. *Brain and Language, 155-156*, 36–43. https://doi.org/10.1016/j.bandl.2016.03.002
Mouchetant-Rostaing, Y., Giard, M. H., Delpuech, C., Echallier, J. F., & Pernier, J. (2000). Early signs of visual categorization for biological and non-biological stimuli in humans. *Neuroreport, 11*(11), 2521–2525. https://doi.org/10.1097/00001756-200008030-00035
Niharika, M. K., & Prema Rao, K. S. (2020). Processing syntax: Perspectives on language specificity. *International Journal of Neuroscience, 130*(8), 841–851. https://doi.org/10.1080/00207454.2019.1707818
Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, Article 156869. https://doi.org/10.1155/2011/156869
Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & cognition, 25*(3), 273–285. https://doi.org/10.3758/bf03211283
Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z
Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension. In In. M. D. Rugg, & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 171–215). Oxford University Press.
Pesciarelli, F., Scorolli, C., & Cacciari, C. (2019). Neural correlates of the implicit processing of grammatical and stereotypical gender violations: A masked and unmasked priming study. *Biological Psychology, 146*, Article 107714. https://doi.org/10.1016/j.biopsycho.2019.06.002
Porkert, J., Siyanova-Chanturia, A., Loerts, H., Schüppert, A., & Keijzer, M. (2024). N400 or P600?—A systematic review of ERP studies on gender stereotype violations (Article) *Language and Linguistics Compass, 18*(5), Article e12530. https://doi.org/10.1111/lnc3.12530.
Proverbio, A. M., Alberio, A., & De Benedetto, F. (2018). Neural correlates of automatic beliefs about gender stereotypes: Males are more prejudicial. *Brain and Language, 186*, 8–16. https://doi.org/10.1016/j.bandl.2018.08.006
Proverbio, A. M., Orlandi, A., & Bianchi, E. (2017). Electrophysiological markers of prejudice related to sexual gender. *Neuroscience, 358*, 1–12. https://doi.org/10.1016/j.neuroscience.2017.06.028
Reali, C., Esaulova, Y., Öttl, A., & von Stockhausen, L. (2015). Role descriptions induce gender mismatch effects in eye movements during reading. *Frontiers in Psychology, 6*, 1607. https://doi.org/10.3389/fpsyg.2015.01607
Rodríguez-Gómez, P., Romero-Ferreiro, V., Pozo, M. A., Hinojosa, J. A., & Moreno, E. M. (2020). Facing stereotypes: ERP responses to male and female faces after gender-stereotyped statements. *Social Cognitive and Affective Neuroscience, 15*(9), 928–940. https://doi.org/10.1093/scan/nsaa117
Sato, S., Casaponsa, A., & Athanasopoulos, P. (2020). Flexing Gender Perception: Brain Potentials Reveal the Cognitive Permeability of Grammatical Information. *Cognitive Science, 44*(9), Article e12884. https://doi.org/10.1111/cogs.12884
Serafini, L., & Pesciarelli, F. (2025). The influence of gender stereotypical primes on the neural processing of words and faces. *Social Cognitive and Affective Neuroscience, 20* (1), Article nsaf031. https://doi.org/10.1093/scan/nsaf031
Siyanova-Chanturia, A., Pesciarelli, F., & Cacciari, C. (2012). The electrophysiological underpinnings of processing gender stereotypes in language. *PloS One, 7*(12), Article e48712. https://doi.org/10.1371/journal.pone.0048712
Su, J.-J., Molinaro, N., Gillon-Dowens, M., Tsai, P.-S., Wu, D. H., & Carreiras, M. (2016). When "he" can also be "she": An ERP study of reflexive pronoun resolution in written Mandarin Chinese. Article 151. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00151
Sutton, S., & Ruchkin, D. S. (1984). The late positive complex. *Advances and New problems Annals of the New York Academy of Sciences, 425*, 1–23. https://doi.org/10.1111/j.1749-6632.1984.tb23520.x
Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience, 2011*, Article 879716. https://doi.org/10.1155/2011/879716
Thom, N., Knight, J., Dishman, R., Sabatinelli, D., Johnson, D. C., & Clementz, B. (2014). Emotional scenes elicit more pronounced self-reported emotional experience and greater EPN and LPP modulation when compared to emotional faces. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 849–860. https://doi.org/10.3758/s13415-013-0225-z
Tremmel, M., & Wahl, I. (2023). Gender stereotypes in leadership: Analyzing the content and evaluation of stereotypes about typical, male, and female leaders. *Frontiers in Psychology, 14*, Article 1034258. https://doi.org/10.3389/fpsyg.2023.1034258
Volpert-Esmond, H. I., Merkle, E. C., & Bartholow, B. D. (2017). The iterative nature of person construal: Evidence from event-related potentials. *Social Cognitive and Affective Neuroscience, 12*(7), 1097–1107. https://doi.org/10.1093/scan/nsx048
Wang, P., Yang, Y. P., Tan, C. H., Zhao, X. X., Liu, Y. H., & Lin, C. D. (2016). Stereotype activation is unintentional: Behavioural and event-related potenials evidence. *International Journal of psychology: Journal International Ďelelőtt psychologie, 51*(2), 156–162. https://doi.org/10.1002/ijop.12135
Wang, P., Yang, Y. P., Tan, C. H., Chen, Q. W., & Cantfort, T. (2017). Gender Stereotype Activation versus Lexical Semantic Activation: An ERP Study. *The Journal of General Psychology, 144*(4), 283–308. https://doi.org/10.1080/00221309.2017.1310714
Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: the time-course of neural activity elicited by specific picture content. *Emotion (Washington, D C), 10*(6), 767–782. https://doi.org/10.1037/a0020242
Weinstein, F., Contreras, A., López, V., & Rodríguez-Gómez, P. (2025). Processing of Descriptive and Prescriptive Stereotypes About Men and Women and Electrophysiological Correlates. *Archives of Sexual Behavior, 54*(5), 1837–1858. https://doi.org/10.1007/s10508-025-03150-1
White, K. R., Crites, S. L., Jr, Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience, 4*(2), 191–198. https://doi.org/10.1093/scan/nsp004
White, K. R. G., Danek, R. H., Herring, D. R., Taylor, J. H., & Crites, S. L. (2018). Taking priming to task: Variations in stereotype priming effects across participant task. *Social Psychology, 49*(1), 29–46. https://doi.org/10.1027/1864-9335/a000326
World Bank Women, business and the law 2023 https://doi.org/10.1596/978-1-4648-1944-52023.
Zhang, X., Li, Q., Sun, S., & Zuo, B. (2018). The time course from gender categorization to gender-stereotype activation. *Social Neuroscience, 13*(1), 52–60. https://doi.org/10.1080/17470919.2016.1251965
Zhang, X., Li, Q., Sun, S., & Zuo, B. (2019). Facial expressions can inhibit the activation of gender stereotypes. *Cognition & Emotion, 33*(7), 1424–1435. https://doi.org/10.1080/02699931.2019.1586648
Zhu, W. L., Fang, P., Xing, H. L., Ma, Y., & Yao, M. L. (2020). Not only top-down: The dual-processing of gender-emotion stereotypes. *Frontiers in Psychology, 11*, 1042. https://doi.org/10.3389/fpsyg.2020.01042