The Effect of Cognitive Control on Gender Stereotype Expression as a Function of Conflict Level

Authors: Chen Li, Shi Xiaoke, Li Weina, Hu Yan

Date: 2022-01-26T00:00:00+00:00

Abstract

With societal development, although gender role division within families has undergone certain changes, the traditional stereotypical notion of "men work outside, women care for the home" still persists. While numerous studies have examined the consistency effect in gender stereotype activation, the influence of cognitive control induced by different conditions on gender stereotype expression under task contexts varying in conflict level remains unclear. Therefore, building upon previous research and grounded in classic dual control theory and conflict monitoring theory, this study investigates the impact of different task conflict trial ratio conditions on gender stereotype expression. The results revealed: (1) When target information was presented as image stimuli, task levels with different conflict ratios showed no difference in gender stereotype expression; (2) When target information was presented as semantic stimuli, task levels with high conflict ratios tended to inhibit gender stereotype expression, whereas task levels with low conflict ratios tended to enhance gender stereotype expression, and the cognitive control processing generated under high conflict levels could be maintained in subsequent different tasks. These findings indicate that cognitive control triggered by tasks with different conflict ratio levels can inhibit gender stereotype expression, but this effect is influenced by information presentation format.

Full Text

The Influence of Cognitive Control Based on Different Conflict Levels on the Expression of Gender Stereotypes

Chen Li#, Xiao-ke Shi#, Wei-na Li#, Yan Hu[1]
[1]School of Psychology, Northwest Normal University, Gansu Provincial Key Laboratory of Behavior and Mental Health, Lanzhou 730070, China

Abstract

With societal development, the division of gender roles in families has undergone changes, yet the traditional stereotype of "men work outside, women work inside" persists. Although numerous studies have examined the consistency effect in gender stereotype activation, the influence of cognitive control induced under different conditions on gender stereotype expression remains unclear across task contexts varying in conflict levels. Building upon previous research, the present study investigated how different task conflict trial ratio conditions affect gender stereotype expression, drawing on classic dual mechanisms of cognitive control theory and conflict monitoring theory. Results revealed: (1) When target information was presented as image stimuli, task conflict ratio levels showed no differential effect on gender stereotype expression; (2) When target information was semantic stimuli, high-conflict ratio tasks inhibited gender stereotype expression, whereas low-conflict ratio tasks enhanced it. Moreover, cognitive control processes generated under high-conflict levels could be maintained across subsequent different tasks. These findings indicate that cognitive control triggered by tasks with varying conflict proportions can inhibit gender stereotype expression, though this effect is moderated by information presentation format.

Keywords: gender stereotype, conflict processing, cognitive control strategy, proactive control, reactive control

Gender stereotypes represent a rigid processing mode of cognition about the two sexes (Rudman & Kilianski, 2000), typically providing socially acceptable standards that communicate which behaviors are appropriate for men and which for women (Croft, 2016). Since ancient times, China has held the belief that "men work outside, women work inside." International research similarly demonstrates that people widely endorse gender stereotypes depicting men as breadwinners and women as caregivers. Such stereotypes foster divergent expectations for gender roles and differential emphasis on skill development, thereby promoting differentiated social behaviors between men and women (Eagly & Steffen, 1984; Eagly, 1987).

Traditional gender stereotypes create considerable conflict for modern women in balancing career and family, while men also bear immense pressure as primary breadwinners. This stereotypical division of domestic roles often negatively impacts gender equality and family harmony (Cundiff & Vescio, 2016). Although macro-level social interventions such as paternity leave, parental leave, and social norm utilization can alleviate role divisions and mitigate public endorsement of gender stereotypes, deeply ingrained gender stereotypes at the individual cognitive level remain uneliminated and can be activated anytime and anywhere. Preventing the automatic activation of stereotypes is a difficult and complex task; however, automatically activated stereotypes can be effectively suppressed by adjusting individual cognitive control (Amodio & Swencionis, 2018).

Cognitive control refers to the ability to make appropriate responses according to goals or tasks, maintain goals in the face of distractors, and suppress habitual or impulsive behaviors (Miller & Cohen, 2001). The Dual Mechanisms of Control (DMC) theory posits that individuals' cognitive control capacity is variable, proposing that cognitive control comprises two primary modes: proactive control and reactive control (Braver et al., 2007). Proactive control emphasizes sustained attention to goal-relevant information throughout a cognitive episode, representing a goal-driven approach to information processing that implements more enduring strategic optimizations in response to rapidly changing external conditions (Funes et al., 2010; Braver et al., 2009). This mode typically reflects a longer macro-timescale developmental pattern (Braver et al., 2007; Miller & Cohen, 2001) and is also termed "focus mode" (Gratton et al., 1992). Reactive control emphasizes perceptual judgments of target stimuli within single cognitive events (Brown et al., 2007), enabling immediate action in response to rapidly changing external conditions. This mode typically represents transient, local, or micro-level control, also known as a late correction mechanism (Purmann et al., 2011).

For example, an individual who habitually turns left at an intersection after work might proactively remind themselves to turn right before reaching the intersection upon hearing about traffic congestion—this illustrates proactive control. Alternatively, they might adopt a reactive control mode, waiting until reaching the intersection, encountering traffic, and then turning right to avoid congestion. In terms of onset and effectiveness, reactive control relies on problem detection and resolution following interference (Braver, 2012), adjusting cognitive patterns after bias emerges. In contrast, proactive control depends on anticipation and prevention before interference occurs, determining cognitive patterns based on goals prior to bias activation (Braver et al., 2009).

Traditionally, theoretical models of bias control in social cognition have emphasized a reactive form of control—cognitive processing implemented after detecting bias to correct its influence on target responses (Fazio, 1990; Amodio et al., 2004; Sherman et al., 2008). Unlike reactive control, proactive control can effectively shift attention from stereotype-related information to the task goal itself before stereotypes are activated, thereby limiting the likelihood that implicitly activated stereotypes will be expressed behaviorally through experiences of task difficulty or increased motivational constraints on expected behavior (Amodio et al., 2018), rather than merely reacting after stereotypes emerge or become activated.

Nearly all studies examining stereotypes inevitably compare processing differences between stereotype-congruent and incongruent information (贾磊 et al., 2016; Lai et al., 2016), with conflict tasks commonly employed in laboratory settings to investigate cognitive control (Gonthier et al., 2021). Conflict monitoring theory (Veen & Carter, 2005) represents the most prevalent theoretical framework for understanding conflict tasks. This theory posits that processing conflict information comprises two distinct components: (1) an evaluative component (conflict monitoring), and (2) an executive component (executive control). The intensity of conflict generated by incongruent information during cognitive processing serves as the foundation for modulating cognitive processing, evoking task-relevant perceptual load. The evaluative component identifies conflict in incoming information and determines whether rapid response is required, while the executive component implements decisions based on task instructions under limited perceptual load (崔诣晨, 2016).

According to dual control theory, individuals flexibly employ reactive or proactive control across different contexts, thereby influencing stereotype expression and intensity. In low-conflict contexts, tasks are relatively simple, stereotype conflict monitoring occurs infrequently, and executive control levels decrease accordingly, making it difficult to suppress stereotype expression. Conversely, in high-conflict contexts, tasks are difficult, conflict monitoring occurs continuously, executive control levels increase appropriately, attention to task goals is enhanced, proactive control mode is facilitated, and stereotype expression is further reduced (Amodio & Swencionis, 2018).

Bartholow and Dickter (2008, 2010) manipulated the ratios of high-conflict and low-conflict trials in racial categorization tasks, revealing that high-conflict groups exhibited a weakened stereotype conflict pattern relative to baseline, whereas low-conflict groups showed a stronger stereotype conflict pattern (Bartholow & Dickter, 2008; Dickter & Bartholow, 2010). However, these studies did not directly verify the influence of cognitive control on racial stereotype expression. In the study by Amodio and Swencionis (2018), the Proactive Model of Control was first proposed (Figure 1 [Figure 1: see original paper]). Using the weapons identification paradigm (Payne, 2001) and stereotype priming tasks, they examined proactive control' s inhibitory effect on racial stereotype expression, confirming that individuals adopted proactive control processing modes under high-conflict levels to suppress racial stereotype expression, whereas they exhibited clear racial stereotypes under low-conflict levels.

Previous research independently investigated the effects of reactive or proactive control on stereotype expression across different task contexts. Moreover, the racial stereotypes examined in these studies represent a cultural phenomenon specific to Western societies, thus failing to provide sufficient evidence for dual mechanisms of control theory. Consequently, the present study focuses on gender stereotypes, which possess greater cultural universality, to investigate the mechanisms through which cognitive control influences stereotype expression within a single experimental task—that is, to identify the conditions under which reactive and proactive control operate in stereotype expression.

The expression of gender stereotypes is related to the format in which gender information is presented. Specifically, gender information presentation formats primarily include two types: pictorial and verbal. Wang et al. (2010) proposed that using concrete pictorial information as stimuli, regardless of how tightly the selected images are stereotypically linked to a particular group (e.g., Black faces-crime images, female faces-housework images), can only reflect the dynamic schematic characteristics of stereotypes but cannot fully capture the abstract processing features implied by the semantic schematic structure of stereotypes. Similarly, using abstract semantic text information as target stimuli cannot fully represent the concrete exemplar characteristics of stereotypes.

Research on context-specific prime effects (CSPC) demonstrates that task contexts with different conflict proportions can elicit varying degrees of attention to conflict or non-conflict information and continuously activate different information processing modes based on conflict specificity (Bailey et al., 2010). In such research, establishing different conflict proportion contexts enables perceivers to better regulate whether stereotypical information undergoes conceptual or perceptual encoding (袁菲, 2015; 崔诣晨, 2016). Furthermore, semantic connections in stereotypes manifest as top-down, conceptually driven processing, whereas concrete exemplar connections exhibit bottom-up, perceptually driven processing (Blair & Banaji, 1996). Therefore, this study further investigates whether the influence of cognitive control induced by different conflict levels on gender stereotype expression demonstrates consistency across information presentation formats.

In real-life situations, beyond enabling cognitive control strategies to function under specific conditions, it is more crucial that cognitive control induced in the current context can create a control preparatory state for subsequent situations (Kleiman et al., 2014), allowing cognitive control to be adjusted promptly across different tasks. Previous research has generally maintained that cognitive control adjustments are domain-specific, with limited control transfer between different domains (Egner, 2008; Hazeltine et al., 2011). However, some perspectives suggest that conflict adjustment states triggered by cognitive control possess domain-general characteristics (Kleiman et al., 2014; 袁菲, 2015). Studies on stereotypes indicate that control preparatory states from classic Flanker tasks in preceding trials can regulate expression in subsequent stereotype Flanker tasks (Kleiman et al., 2014; 袁菲, 2015). Although existing research has attempted to demonstrate cross-task consistency in cognitive control' s effect on stereotype expression, this question remains inadequately addressed because identical task paradigms were used across consecutive experimental trials.

In summary, this study investigates the mechanisms through which cognitive control influences gender stereotype expression. Addressing the aforementioned issues, the research hypotheses are: (1) Target classification tasks at high conflict levels will elicit proactive control, thereby inhibiting gender stereotype expression; target classification tasks at low conflict levels will elicit reactive control, thereby facilitating gender stereotype expression. (2) The inhibitory effect of cognitive control induced by different conflict levels on gender stereotype expression demonstrates cross-task consistency.

Experiment 1

Participants Sample size was calculated using G*Power 3.1 (Faul et al., 2009), with effect size set at 0.25, α at 0.05, and power at 0.80, indicating a minimum requirement of 36 participants. Based on sample sizes used in previous similar studies (ranging from 25 to 63; Amodio & Swencionis, 2018; Kleiman et al., 2014) and considering potential issues during experimentation (e.g., invalid participants, gender imbalance), we recruited 60 participants (26 male) with a mean age of 24.60 years (SE = 2.82). All participants had normal or corrected-to-normal vision, were proficient in computer operation, and had not previously participated in similar experiments. Participation was voluntary, and participants received compensation upon completion.

Materials Experimental materials consisted of "prime-target" stimulus pairs, using male and female face pictures as primes and work and housework scene pictures as targets. Prime pictures included four male and four female images, with two male and two female images used as practice stimuli and the remaining two male and two female images used as formal experimental stimuli. Target pictures comprised six housework images and six work images, with two housework and two work images serving as practice stimuli and four housework and four work images serving as formal experimental stimuli.

Prime pictures were selected following Croft's (2016) Study 1, with ten male and ten female images chosen from existing databases and standardized using Photoshop 7.0 software to ensure uniform gray background brightness and dimensions of 420$×$380 pixels (Amodio & Swencionis, 2018). An independent sample of 176 participants rated the 20 selected images on a 5-point scale for overall facial impression (1 = "very negative" to 5 = "very positive"). Eight male and eight female images with intermediate overall impression ratings were selected. Difference tests on the selected images revealed no significant gender differences in overall impression (Mmale = 3.10, Mfemale = 3.20), t(14) = -1.494, p = 0.157. Additionally, research indicates that people hold different stereotypes for men and women of different ages (任娜 et al., 2012); therefore, this study controlled face age between 25-30 years, with no significant age difference between male and female images (Mmale = 27.95, Mfemale = 27.37), t(14) = 0.492, p = 0.631. Furthermore, one-sample t-tests confirmed that both male and female images differed significantly from extreme overall impression levels (ps $<$ 0.001).

Target picture materials were sourced from work and family images used in Croft' s (2016) research. Twelve scene images related to housework activities and twelve related to professional activities were selected from online databases (containing no people, only relevant work or home backgrounds). Processing standards matched those for prime pictures, ensuring consistent brightness and dimensions of 420$×$380 pixels (Croft, 2016). An independent sample of 176 participants rated the 24 selected images on a 5-point scale for scene representativeness (1 = "highly unrepresentative" to 5 = "highly representative"). Six images representing housework scenes and six representing work scenes were selected. Difference tests revealed no significant difference in representativeness between work and housework images (Mwork = 3.79, Mhousework = 3.82), t(10) = 0.295, p = 0.779. One-sample t-tests confirmed that both work and housework images differed significantly from the midpoint of the representativeness scale (ps < 0.001). Finally, selected images were converted to BMP format using Adobe Photoshop CS1.

Experimental Design A 3 (task conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture vs. female picture) × 2 (target stimulus: work picture vs. housework picture) within-subjects design was employed. This yielded 12 experimental conditions: for each conflict level, male picture-work picture and female picture-housework picture (i.e., target pictures congruent with stereotypes activated by primes) and male picture-housework picture and female picture-work picture (i.e., target pictures incongruent with stereotypes activated by primes). Dependent variables were mean response time and accuracy for each experimental condition in the picture classification task.

Procedure Given that individuals' initial emotional states and gender role ideology levels can influence stereotype cognitive processing, the revised Positive and Negative Affect Schedule (PANAS; 邱林 et al., 2008) was administered to measure participants' emotional states. This 5-point scale includes positive affect self-ratings (nine positive adjectives) and negative affect self-ratings (nine negative adjectives), with higher scores indicating stronger emotions. The Gender Role Ideology Scale was developed by integrating items from the Chinese Women' s Social Status Survey (CWSSS) and Chinese General Social Survey (CGSS), distinguishing between traditional and modern orientations (Rajadhyaksha et al., 2015). In this experiment, the Gender Role Ideology Scale demonstrated a reliability coefficient of 0.83 and structural validity of 0.67. The 60 participants scored M = 3.80, SD = 0.51 on this scale. Regarding family role division, 53.3% reported father-as-breadwinner/mother-as-homemaker arrangements, 3.3% reported mother-as-breadwinner/father-as-homemaker arrangements, and 43.3% reported balanced parental roles.

The experimental flow was controlled using a modified weapons identification task (Payne, 2001) and methods for manipulating task interference (Appelbaum et al., 2014). The experiment comprised three conditions: high-conflict, baseline, and low-conflict. In the high-conflict condition, incongruent trials were presented at 80% and congruent trials at 20%; in the baseline condition, both incongruent and congruent trials were presented at 50%; in the low-conflict condition, incongruent trials were presented at 20% and congruent trials at 80%. The formal experiment consisted of six blocks, with two blocks per conflict level, arranged using a balanced Latin square design. Each block contained 40 trials, totaling 240 trials. Specifically, the high-conflict level included 16 congruent and 64 incongruent trials; the baseline level included 40 congruent and 40 incongruent trials; the low-conflict level included 64 congruent and 16 incongruent trials. Stimuli were presented using E-Prime 2.0, with all stimuli appearing at the center of the computer screen and trial order randomized within each block.

Experimental instructions were: "First, a '+' fixation point will appear at the center of the screen; please focus your attention on it. Subsequently, a gender face picture will be presented at the center, followed by a scene picture, and finally an oval-shaped picture. When the oval-shaped picture appears, you must make a rapid and accurate categorical judgment about the scene picture: press the D key for housework images and the K key for work images. Press the spacebar to begin the practice phase when you understand the instructions." Practice stimuli differed from formal experimental materials, with 16 practice trials total. If participants' accuracy fell below 70%, they were required to repeat practice until reaching the accuracy criterion before proceeding to the formal experiment. A single trial procedure is illustrated in Figure 2 [Figure 2: see original paper].

Results

A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work picture, housework picture) repeated measures ANOVA on mean accuracy across conflict conditions revealed a significant interaction between prime and target stimuli, $F(1, 59) = 4.91$, $p = 0.031$, $p^2 = 0.08$. No other main effects or interactions were significant ($ps > 0.05$).

A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work picture, housework picture) repeated measures ANOVA on mean response times revealed no significant effects ($ps > 0.05$).

When gender role ideology scores were entered as covariates in repeated measures ANOVAs on mean accuracy and response time across conflict levels, the main effect of conflict level remained non-significant ($ps > 0.05$).

Table 1 presents accuracy and response times for target classification under different conflict levels and target stimuli. Table 2 presents proactive and reactive control estimates across different conflict levels.

Discussion

Although our hypotheses were not supported by the analytical results, a closer examination of the implications underlying response time and accuracy measures (Kidder et al., 2018) reveals that despite non-significant differences in accuracy across the three conflict levels, both accuracy and control estimates remained relatively high (see Table 2). This suggests that participants' cognitive accuracy for target stimuli was unaffected by conflict level interference throughout the task. The absence of significant response time differences across conflict levels indicates that participants' processing speed for target stimuli was also unaffected by cognitive conflict interference, remaining consistently stable at the same attentional level.

Through literature review and consideration of practical factors, we preliminarily attribute the failure to validate our hypotheses to several aspects. First, regarding task nature, categorical judgments based on presented target pictures generate different cognitive conflicts than other stereotype measurement tasks (e.g., prime-target congruency judgment tasks, Implicit Association Test). The connectionist model of stereotype representation posits that distributed recurrent networks excel at identifying image or pattern interference, with recurrent connections emerging based on specific tasks enabling individuals to essentially filter out various interferences (Van et al., 2003). Consequently, not emphasizing the connection judgment between prime/distractor stimuli and target pictures makes it easier for individuals to adopt a processing mode focusing solely on target pictures, with corresponding judgments being less influenced by differences in conflict level within the same conflict type.

Second, regarding the external characteristics of experimental materials, both prime and target stimuli were colorful, high-resolution pictures with high ecological validity. However, the numerous repeated pairings of prime and target stimuli likely induced practice effects across different conflict task levels after participants learned the material processing mode. Moreover, extensive research on priming stimuli has found that material richness critically impacts accuracy and response time (Kidder et al., 2018).

Furthermore, considering the internal processing format of experimental materials, theoretical perspectives on differences between text and image processing suggest that text processing involves more additional processing than image processing before entering emotional judgment. This additional text processing includes top-down processing that generates mental representations helping us access emotional aspects of stimuli. Due to self-protective instincts, negative mental representations of text may be weakened through individuals' conscious systems. Consequently, emotional biases may differ between text and pictures (Yuan et al., 2019).

Experiment 2

Participants As in Experiment 1, the minimum required sample size was 36. We recruited 62 participants, of whom two were excluded for failing to follow task instructions, three were excluded due to excessive invalid responses, and two were excluded for low accuracy. The final valid sample comprised 55 participants (26 male) with a mean age of 19.87 years (SD = 2.54). All participants had normal or corrected-to-normal vision, were proficient in computer operation, and had not previously participated in similar experiments. Participation was voluntary, and participants received compensation upon completion.

Experimental Design A 3 (task conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work word, housework word) within-subjects design was employed. This yielded 12 experimental conditions: for each conflict level, male picture-work word and female picture-housework word (i.e., target words congruent with stereotypes activated by primes) and male picture-housework word and female picture-work word (i.e., target words incongruent with stereotypes activated by primes). Dependent variables were mean response time and accuracy for each experimental condition in the word classification task.

Materials Experimental materials consisted of "prime-target" stimulus pairs, using male and female face pictures as primes and work and housework words as targets. Prime pictures included six male and six female images, with two male and two female images used as practice stimuli and four male and four female images used as formal experimental stimuli. Target words comprised ten housework activity words and ten professional work words, with two housework and two work words serving as practice stimuli and eight housework and eight work words serving as formal experimental stimuli.

Prime pictures were selected from the gender pictures screened in Experiment 1. Target words were sourced from 吴梦玲' s (2017) research and supplemented through online surveys to obtain 20 housework-related and 20 profession-related words. Forty participants (19 male, M = 20 years) rated word typicality on a 5-point scale (1 = "very atypical" to 5 = "very typical"). Difference tests revealed no significant typicality difference between housework and work words (Mhousework = 3.95, Mwork = 3.97), t(38) = -0.288, p = 0.775. One-sample t-tests confirmed that both housework and work words differed significantly from the midpoint of the typicality scale (ps < 0.001). Words with the highest mean scores in each category were selected as target words.

Procedure Questionnaires measuring participants' emotional states and gender role ideology levels were identical to Experiment 1. In this experiment, the Gender Role Ideology Scale demonstrated a reliability coefficient of 0.79 and structural validity of 0.62. The 55 participants scored M = 3.89, SD = 0.49 on this scale. Regarding family role division, 47.3% reported father-as-breadwinner/mother-as-homemaker arrangements, 0% reported mother-as-breadwinner/father-as-homemaker arrangements, 49.1% reported balanced parental roles, and 3.6% reported neither parent fulfilling both roles. The overall experimental flow was consistent with Experiment 1, with the difference that Experiment 2 required categorical judgments of target words as "housework" or "profession." The specific trial procedure is illustrated in Figure 3 [Figure 3: see original paper].

Results

A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work word, housework word) repeated measures ANOVA on mean accuracy across conflict conditions revealed significant main effects of conflict level, $F(2, 108) = 7.94$, $p = 0.001$, $p^2 = 0.13$, and target stimulus, $F(1, 54) = 9.69$, $p = 0.003$, $p^2 = 0.15$, along with a significant prime × target interaction, $F(1, 54) = 16.37$, $p < 0.001$, $p^2 = 0.23$.

Critically, the three-way interaction among conflict level, prime, and target was also significant, $F(2, 108) = 3.65$, $p = 0.029$, $p^2 = 0.06$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of conflict level on accuracy (Table 3) revealed that mean accuracy for the four stimulus types was higher in the high-conflict condition (Mhigh-conflict = 0.96, SD = 0.01) than in the low-conflict condition (Mlow-conflict = 0.94, SD = 0.01, $p < 0.001$) and baseline condition (Mbaseline = 0.95, SD = 0.01, $p = 0.019$). However, the difference between low-conflict (Mlow-conflict = 0.94, SD = 0.01) and baseline conditions (Mbaseline = 0.95, SD = 0.01) was not significant ($p = 0.228$). Thus, as predicted, participants performed best under high-conflict levels.

To decompose the three-way interaction effect on accuracy, we examined the prime × target interaction separately at each conflict level through 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work word, housework word) repeated measures ANOVAs. Results are illustrated in Figure 4 [Figure 4: see original paper]. First, under baseline conditions, neither main effect was significant ($ps > 0.05$), but the prime × target interaction was significant, $F(1, 54) = 7.77$, $p = 0.007$, $p^2 = 0.13$. Simple effects analysis revealed a typical stereotype pattern: following male pictures, responses to work words (M = 0.98, SD = 0.01) were more accurate than to housework words (M = 0.93, SD = 0.01), $F(1, 54) = 11.28$, $p = 0.001$, $p^2 = 0.17$. In contrast, following female pictures, accuracy for work words (M = 0.94, SD = 0.01) and housework words (M = 0.95, SD = 0.01) did not differ, $F(1, 54) = 0.04$, $p = 0.836$.

As expected, responses in the low-conflict condition showed a larger stereotype effect than baseline: the prime × target interaction was significant, $F(1, 54) = 12.28$, $p = 0.001$, $p^2 = 0.19$, with no significant main effects ($ps > 0.05$). Simple effects analysis showed that following male pictures, responses to work words (M = 0.97, SD = 0.01) were more accurate than to housework words (M = 0.92, SD = 0.01), $F(1, 54) = 18.02$, $p < 0.001$, $p^2 = 0.25$, whereas following female pictures, accuracy for work words (M = 0.93, SD = 0.01) and housework words (M = 0.95, SD = 0.01) did not differ, $F(1, 54) = 1.14$, $p = 0.291$.

In contrast, under high-conflict conditions, stereotypes were eliminated: no significant prime × target interaction emerged, $F(1, 54) = 0.21$, $p = 0.651$, and no significant main effects were observed ($ps > 0.05$). Accuracy for work words following male pictures (M = 0.97, SD = 0.01) did not differ from housework words following male pictures (M = 0.95, SD = 0.01), $F(1, 54) = 3.40$, $p = 0.071$, nor did accuracy for work words following female pictures (M = 0.97, SD = 0.01) differ from housework words following female pictures (M = 0.96, SD = 0.01), $F(1, 54) = 1.11$, $p = 0.298$.

Subsequent repeated measures ANOVA on mean response times across conflict conditions revealed similar patterns: a significant main effect of conflict level, $F(2, 108) = 5.99$, $p = 0.003$, $p^2 = 0.10$, and a significant prime × target interaction, $F(1, 54) = 8.54$, $p = 0.005$, $p^2 = 0.14$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of conflict level on response time (Table 3) showed that mean response time for the four stimulus types was significantly longer in the high-conflict condition (Mhigh-conflict = 467.78, SD = 9.77) than in the low-conflict condition (Mlow-conflict = 442.45, SD = 10.17, $p < 0.001$). The baseline condition (Mbaseline = 456.70, SD = 8.13) marginally differed from the low-conflict condition (Mlow-conflict = 442.45, SD = 10.17, $p = 0.076$). However, response times in the high-conflict condition (Mhigh-conflict = 467.78, SD = 9.77) did not differ significantly from baseline (Mbaseline = 456.70, SD = 8.15, $p = 0.150$). Thus, as predicted, participants showed longest response times under high-conflict levels and shortest under low-conflict levels.

When gender role ideology scores were entered as covariates in repeated measures ANOVAs on mean accuracy and response time across conflict levels, the main effect of conflict level remained significant for accuracy, $F(2, 106) = 4.99$, $p = 0.014$, $p^2 = 0.09$, and for response time, $F(2, 106) = 3.31$, $p = 0.04$, $p^2 = 0.06$.

We further tested our hypotheses using the Process Dissociation Procedure (PDP; Jacoby, 1991; Jacoby et al., 1999). PDP is a method for assessing the unique contributions of controlled processing (i.e., task goal-congruent) and automatic processing (i.e., stereotype bias-congruent) to task performance from behavioral patterns. Within the PDP framework, the control estimate (PDP-C) represents the probability that an individual will respond in an accurate, goal-congruent manner without stereotype-driven bias from gender priming (P[correct response on congruent trials] - P[error on incongruent trials]). The automaticity estimate (PDP-A) represents the probability that, to the extent control fails, an individual's response will be influenced by gender-primed stereotypes due to stereotypical associations with the target (P[stereotypical error/(1 - control)]; Payne, 2001).

A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) repeated measures ANOVA on PDP-C estimates revealed a significant main effect of conflict level, $F(2, 108) = 7.94$, $p = 0.001$, $p^2 = 0.13$, but no significant main effect of prime stimulus or interaction ($ps > 0.05$). Thus, control processing estimates across different conflict levels did not differ between male and female prime picture types.

Further multiple comparisons indicated that the control processing estimate was significantly higher in the high-conflict condition (M = 0.93, SD = 0.01) than in the low-conflict condition (M = 0.88, SD = 0.01, $p < 0.001$) and baseline condition (M = 0.90, SD = 0.01, $p = 0.019$). However, no significant difference emerged between baseline and low-conflict conditions ($p = 0.228$; see Table 4). High-conflict levels elicited the highest degree of control processing, low-conflict levels elicited the least, and baseline control processing estimates fell in between. This analysis essentially mirrors the previously reported main effect of conflict level on accuracy, as PDP-C represents mean accuracy as a function of conflict level and prime stimulus. This pattern further supports our hypothesis that high-conflict processing tasks more readily elicit proactive control.

Finally, we analyzed PDP-A estimates to test our predictions. Because the PDP-A formula cannot be solved when PDP-C = 1 (i.e., when accuracy for one or more trial types reaches 100%), some participants had missing PDP-A values. Only 20 participants had valid PDP-A values across all trial conditions and were included in this analysis. A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) repeated measures ANOVA on these estimates revealed no significant main effects or interactions ($ps > 0.05$). However, data showed that PDP-A estimates were highest in the high-conflict condition ($M = 0.69$, $SD = 0.05$), intermediate at baseline ($M = 0.63$, $SD = 0.05$), and lowest in the low-conflict condition ($M = 0.56$, $SD = 0.06$; see Table 4). This pattern is consistent with conflict effects modulating control processing engagement without affecting implicit associations (Hilgard et al., 2015).

Discussion

Results from word classification task accuracy provided strong evidence for the control processing advantage induced by high interference. Specifically, proactive control facilitated processing of global task goals in high-conflict ratio tasks by shifting attention away from task-irrelevant distractors (e.g., gender; Amodio & Swencionis, 2018). Under high-conflict conditions, participants made accurate judgments for both housework- and work-related words regardless of whether male or female pictures served as primes. In contrast, baseline and low-conflict conditions exhibited typical gender stereotype patterns: when male pictures served as primes, participants showed more accurate judgments for work-related words, whereas when female pictures served as primes, they showed more accurate judgments for housework-related words. Additionally, response times were longer under high-conflict and baseline levels than under low-conflict conditions. PDP analyses of control processing across conflict levels revealed higher control processing estimates for high-conflict and baseline levels and lower estimates for low-conflict levels. We therefore speculate that low-conflict levels may increase participants' reliance on reactive control, whereas high-conflict levels promote greater dependence on proactive control. However, dual mechanisms of cognitive control theory posits that reactive and proactive control processes can change dynamically according to circumstances (Braver, 2012; Schmid et al., 2015). The current finding that automatic processing estimates were higher in high-conflict and baseline conditions while control processing estimates were lower in low-conflict conditions indirectly demonstrates the cooperative processing nature of reactive and proactive control within the same task (Funes et al., 2010).

In this study, our survey of family role division revealed that in most families, "fathers tend more toward breadwinning while mothers tend more toward caregiving." This role division difference aligns more closely with gender stereotype presentation in low-conflict tasks. In high-conflict tasks, however, stereotype-congruent stimuli appeared at very low proportions while conflict stimuli appeared at higher proportions—an anti-stereotype presentation format uncommon in real life. According to the mixed model of stereotypes (Sherman, 1996), text stimuli in low-conflict conditions, due to principles of cognitive economy, tend toward abstract representations, whereas high-conflict and baseline conditions, due to the specificity of paired stimuli, align more with exemplar representations, which are characterized by accuracy and slower processing speed.

Experiment 3

Participants As in Experiments 1 and 2, the minimum required sample size was 36. We recruited 67 participants, of whom four were excluded for failing to complete the entire experimental procedure and five were excluded due to excessive invalid responses. The final valid sample comprised 58 participants (29 male) with a mean age of 22.33 years (SD = 2.67). All participants had normal or corrected-to-normal vision, were proficient in computer operation, and had not previously participated in similar experiments. Participation was voluntary, and participants received compensation upon completion.

Experimental Design A 3 (conflict level in word classification task: high-conflict, baseline, low-conflict) × 2 (target category in gender Flanker task: masculine name, feminine name) × 2 (flanker category in gender Flanker task: masculine trait word, feminine trait word) within-subjects design was employed, yielding 12 experimental conditions. For each conflict level, masculine trait word-masculine name-masculine trait word and feminine trait word-feminine name-feminine trait word (i.e., target names congruent with stereotypes activated by flankers) and masculine trait word-feminine name-masculine trait word and feminine trait word-masculine name-feminine trait word (i.e., target names incongruent with stereotypes activated by flankers) were included. Dependent variables were mean response time and accuracy for both the word classification task and the gender Flanker task.

Combinations of names and stereotypical trait words created 48 gender stereotype-congruent stimuli (e.g., masculine names surrounded by masculine stereotypical trait words, "strong Li Jun strong" ; including 24 male-congruent and 24 female-congruent pairings) and 48 gender stereotype-incongruent stimuli (e.g., masculine names surrounded by feminine stereotypical trait words, "fragile Li Jun fragile" ; including 24 male-incongruent and 24 female-incongruent pairings). To reduce practice effects from identical stimulus materials across conflict levels, relatively homogeneous but non-identical stimulus materials were used at each conflict level.

Questionnaires measuring participants' emotional states and gender role ideology levels were identical to Experiment 1. In this experiment, the Gender Role Ideology Scale demonstrated internal consistency reliability of 0.84 and structural validity of 0.71. The 58 participants scored M = 3.76, SD = 0.50 on this scale.

Regarding family role division, 51.7% reported father-as-breadwinner/mother-as-homemaker arrangements, 5.2% reported mother-as-breadwinner/father-as-homemaker arrangements, 41.4% reported balanced parental roles, and 1.7% reported neither parent fulfilling both roles.

Materials Experimental materials comprised "prime-target" and "gender trait word-name-gender trait word" pairs, involving four stimulus types: prime pictures, target words, gender trait words, and names. Male and female face pictures served as primes, work and housework words as targets, gender trait words as flankers, and names as targets in the Flanker task. Prime pictures and target words were selected from materials screened in Experiments 1 and 2. Gender trait words and names were selected based on domestic stereotype research (陈莉, 2011; 刘红豆, 2018) and supplemented through online surveys to obtain 60 masculine stereotypical trait words, 60 feminine stereotypical trait words, 40 masculine names, and 40 feminine names. An additional 50 participants (20 male, M = 25.41 years) rated the 120 trait words on a 7-point gender bias scale (1 = extremely masculine, i.e., "highly feminine" ; 2 = "relatively feminine" ; 3 = lower masculinity; 4 = "neutral," applicable to both genders; 5 = "generally feminine" ; 6 = "relatively feminine" ; 7 = "highly feminine") and the 80 names on a 5-point gender bias scale. Forty-eight gender-stereotypical trait words and 52 names showing clear gender bias were selected. Difference tests confirmed significant rating differences between masculine (M = 5.37) and feminine (M = 2.73) trait words, $t(46) = 0.250$, $p < 0.001$, with both differing significantly from neutral ($ps < 0.001$). Masculine (M = 4.62) and feminine (M = 1.37) names also differed significantly, $t(50) = 98.201$, $p < 0.001$, with both differing from neutral ($ps < 0.001$).

Selected stimulus materials were categorized as follows: - Prime pictures: Eight male and eight female pictures, including practice stimuli and two male and two female pictures for each conflict level. - Target words: Fourteen housework activity words and fourteen professional work words, including two housework and two work words as practice stimuli and four housework and four work words for each conflict level. - Gender trait words: Twenty-four masculine and twenty-four feminine trait words, including four masculine and four feminine trait words as practice stimuli and the remainder as formal experimental stimuli. - Names: Twenty-six masculine and twenty-six feminine names, including two masculine and two feminine names as practice stimuli and the remainder as formal experimental stimuli.

Procedure The experiment utilized an improved trial-to-trial control adjustment paradigm (Kleiman et al., 2014) to manipulate experimental flow. Each complete trial comprised two consecutive sub-trials involving two different behavioral response tasks (see Figure 5 [Figure 5: see original paper] for a complete trial procedure). The first task was the word classification task used in Experiment 2. To verify the influence of different conflict levels on subsequent processing tasks, conflict classification standards matched Experiment 2. The second task was a gender Flanker task. In this task, participants saw a masculine name (e.g., Li Jun) or feminine name (e.g., Liu Ting) at the center of the screen, with stereotypical trait words that were gender-congruent or incongruent appearing to the left and right of each name. Participants judged the name' s gender by pressing the D or K key. Participants completed practice trials after understanding the procedure. Practice stimuli did not appear in the formal experiment, with 16 practice trials total. If accuracy fell below 70%, participants repeated practice until reaching the criterion. The formal experiment consisted of three blocks, with word classification tasks and gender Flanker tasks at different conflict levels constituting separate blocks arranged in a balanced Latin square design. Each block contained 80 complete trials, totaling 240 trials. In the high-conflict level, 16 congruent and 64 incongruent classification trials were presented. In the baseline level, 40 congruent and 40 incongruent classification trials were presented. In the low-conflict level, 64 congruent and 16 incongruent classification trials were presented. Conflict and compatible trials in the gender Flanker task appeared randomly across blocks at consistent proportions.

Results

Word Classification Task Results A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work word, housework word) repeated measures ANOVA on mean accuracy revealed significant main effects of conflict level, $F(2, 114) = 6.01$, $p = 0.003$, $p^2 = 0.10$, and a significant prime × target interaction, $F(1, 57) = 13.34$, $p = 0.001$, $p^2 = 0.19$. Critically, the three-way interaction among conflict level, prime, and target was also significant, $F(2, 114) = 7.22$, $p = 0.001$, $p^2 = 0.11$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of conflict level (Table 5) showed that mean accuracy for the four stimulus types was significantly higher in the high-conflict condition (Mhigh-conflict = 0.91, SD = 0.01) than in the low-conflict condition (Mlow-conflict = 0.87, SD = 0.01, $p = 0.002$) and baseline condition (Mbaseline = 0.87, SD = 0.01, $p = 0.009$). However, the difference between low-conflict (Mlow-conflict = 0.87, SD = 0.01) and baseline conditions (Mbaseline = 0.87, SD = 0.01) was not significant ($p = 0.904$).

To decompose the three-way interaction, we conducted separate 2 (prime stimulus: male picture, female picture) × 2 (target stimulus: work word, housework word) ANOVAs for each conflict type. Results are illustrated in Figure 6 [Figure 6: see original paper]. First, under baseline conditions, the prime × target interaction was significant, $F(1, 57) = 11.75$, $p = 0.001$, $p^2 = 0.17$. Simple effects analysis revealed a typical stereotype pattern: when male pictures served as primes, accuracy for housework words (M = 0.85, SD = 0.02) did not differ significantly from work words (M = 0.88, SD = 0.02), $F(1, 57) = 1.12$, $p = 0.293$. However, when female pictures served as primes, accuracy for work words (M = 0.84, SD = 0.02) was lower than for housework words (M = 0.91, SD = 0.01), $F(1, 57) = 8.54$, $p = 0.005$, $p^2 = 0.13$.

As expected, responses in the low-conflict condition showed a larger stereotype effect than baseline: the prime × target interaction was significant, $F(1, 57) = 15.81$, $p < 0.001$, $p^2 = 0.22$. Simple effects analysis revealed that when male pictures served as primes, accuracy for housework words (M = 0.84, SD = 0.02) did not differ significantly from work words (M = 0.89, SD = 0.02), $F(1, 57) = 1.98$, $p = 0.165$. However, when female pictures served as primes, accuracy for work words (M = 0.83, SD = 0.02) was lower than for housework words (M = 0.92, SD = 0.01), $F(1, 57) = 10.78$, $p = 0.002$, $p^2 = 0.16$.

Compared to baseline, the stereotype pattern was attenuated under high-conflict conditions, with no significant prime × target interaction, $F(1, 57) = 0.24$, $p = 0.626$. When male pictures served as primes, accuracy for housework words (M = 0.90, SD = 0.01) did not differ significantly from work words (M = 0.91, SD = 0.02), $F(1, 57) = 0.05$, $p = 0.820$. When female pictures served as primes, accuracy for work words (M = 0.90, SD = 0.01) did not differ significantly from housework words (M = 0.91, SD = 0.02), $F(1, 57) = 0.11$, $p = 0.744$.

Separate repeated measures ANOVAs on mean response times for each conflict condition revealed: a significant main effect of conflict level, $F(2, 114) = 21.10$, $p < 0.001$, $p^2 = 0.27$, and a significant prime × target interaction, $F(1, 57) = 5.70$, $p = 0.02$, $p^2 = 0.10$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of conflict level (Table 6) showed that mean response time for the four stimulus types differed significantly between high-conflict (Mhigh-conflict = 525.33, SD = 15.81) and low-conflict conditions (Mlow-conflict = 465.52, SD = 13.22, $p < 0.001$), and between baseline (Mbaseline = 513.87, SD = 13.10) and low-conflict conditions (Mlow-conflict = 465.52, SD = 13.22, $p < 0.001$). However, high-conflict (Mhigh-conflict = 525.33, SD = 15.81) and baseline conditions (Mbaseline = 513.87, SD = 13.10) did not differ significantly ($p = 0.244$). Participants showed longest response times under high-conflict levels and shortest under low-conflict levels.

When gender role ideology scores were entered as covariates in repeated measures ANOVAs on mean accuracy and response time across conflict levels, the main effect of conflict level remained significant for accuracy, $F(2, 112) = 9.08$, $p = 0.001$, $p^2 = 0.14$, and for response time, $F(2, 112) = 5.78$, $p = 0.004$, $p^2 = 0.09$.

We then used PDP to examine whether tasks under different conflict levels differentially influenced subsequent task processing. A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) repeated measures ANOVA on PDP-C estimates revealed a significant main effect of conflict level, $F(2, 114) = 6.01$, $p = 0.003$, $p^2 = 0.10$, but no significant main effect of prime stimulus or interaction. Thus, control processing estimates across conflict levels did not differ between male and female prime picture types. Further multiple comparisons indicated that the control processing estimate was significantly higher in the high-conflict condition (M = 0.81, SD = 0.02) than in the low-conflict condition (M = 0.74, SD = 0.03, $p = 0.009$) and baseline condition (M = 0.74, SD = 0.03, $p = 0.002$). However, baseline and low-conflict conditions did not differ significantly ($p = 0.904$; see Table 6). High-conflict levels elicited the highest degree of control processing, low-conflict levels elicited the least, and baseline estimates fell in between. This analysis essentially mirrors the previously reported main effect of conflict level on accuracy, as PDP-C represents mean accuracy as a function of conflict level and prime stimulus. This pattern further supports our hypothesis that high-conflict processing tasks more readily elicit proactive control.

Finally, we analyzed PDP-A estimates to test our predictions. Because the PDP-A formula cannot be solved when PDP-C = 1 (i.e., when accuracy for one or more trial types reaches 100%), some participants had missing PDP-A values. Thirty-seven participants had valid PDP-A values across all trial conditions and were included in this analysis. A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) repeated measures ANOVA on PDP-A estimates revealed no significant main effects or interactions ($ps > 0.05$). Automatic processing estimates across conflict levels are presented in Table 6 , consistent with Experiment 2 results.

Gender Flanker Task Results In the gender Flanker task, a 3 (previous task conflict level: high-conflict, baseline, low-conflict) × 2 (target category: masculine name, feminine name) × 2 (trait word category: masculine trait word, feminine trait word) repeated measures ANOVA on accuracy revealed significant main effects of previous task conflict level, $F(2, 56) = 6.58$, $p = 0.003$, $p^2 = 0.19$, and a significant target category × trait word category interaction, $F(1, 57) = 21.87$, $p < 0.001$, $p^2 = 0.28$. Critically, the three-way interaction was also significant, $F(2, 114) = 3.91$, $p = 0.023$, $p^2 = 0.06$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of previous task conflict level (Table 7) showed that mean accuracy for the four stimulus types was significantly higher in the high-conflict condition (Mhigh-conflict = 0.90, SD = 0.01) than in the low-conflict condition (Mlow-conflict = 0.83, SD = 0.02, $p = 0.001$) and baseline condition (Mbaseline = 0.85, SD = 0.02, $p = 0.009$). However, the difference between low-conflict (Mlow-conflict = 0.83, SD = 0.02) and baseline conditions (Mbaseline = 0.85, SD = 0.02) was not significant ($p = 0.061$).

To further examine how tasks at different conflict levels influenced accuracy in subsequent Flanker tasks, we conducted separate 2 (target category: masculine name, feminine name) × 2 (trait word category: masculine trait word, feminine trait word) ANOVAs on accuracy for Flanker tasks following each conflict level. Results are illustrated in Figure 7 [Figure 7: see original paper]. First, under baseline conditions, the target category × trait word category interaction was significant, $F(1, 57) = 21.69$, $p < 0.001$, $p^2 = 0.28$. Simple effects analysis revealed a typical stereotype pattern: when Flanker task trait words were masculine, judgments of masculine names (M = 0.89, SD = 0.02) were more accurate than feminine names ($M = 0.80$, $SD = 0.03$), $F(1, 57) = 22.53$, $p < 0.001$, $p^2 = 0.28$. When trait words were feminine, judgments of masculine names ($M = 0.83$, $SD = 0.02$) were less accurate than feminine names ($M = 0.88$, $SD = 0.02$), $F(1, 57) = 9.19$, $p = 0.004$, $p^2 = 0.14$.

As expected, responses in the low-conflict condition also showed typical stereotype effects: the target category × trait word category interaction was significant, $F(1, 57) = 14.84$, $p < 0.001$, $p^2 = 0.21$. Simple effects analysis revealed that when Flanker task trait words were masculine, correct judgments of masculine names ($M = 0.88$, $SD = 0.02$) exceeded those of feminine names ($M = 0.80$, $SD = 0.03$), $F(1, 57) = 14.29$, $p < 0.001$, $p^2 = 0.20$. When trait words were feminine, correct judgments of masculine names ($M = 0.79$, $SD = 0.03$) were lower than feminine names ($M = 0.85$, $SD = 0.02$), $F(1, 57) = 7.98$, $p = 0.007$, $p^2 = 0.12$.

Compared to baseline, the stereotype pattern was attenuated under high-conflict conditions, with no significant target category × trait word category interaction, $F(1, 57) = 3.68$, $p = 0.060$. When Flanker task trait words were masculine, accuracy for masculine names ($M = 0.91$, $SD = 0.01$) did not differ from feminine names ($M = 0.88$, $SD = 0.02$), $F(1, 57) = 3.28$, $p = 0.075$. When trait words were feminine, accuracy for masculine names ($M = 0.89$, $SD = 0.02$) did not differ from feminine names ($M = 0.91$, $SD = 0.01$), $F(1, 57) = 2.24$, $p = 0.140$.

Repeated measures ANOVA on response times for the gender Flanker task revealed similar patterns: a significant main effect of previous task conflict level, $F(2, 114) = 4.08$, $p = 0.02$, $p^2 = 0.07$, and a significant target category × trait word category interaction, $F(1, 57) = 20.30$, $p < 0.001$, $p^2 = 0.26$. No other main effects or interactions were significant ($ps > 0.05$).

Post-hoc multiple comparisons on the main effect of previous task conflict level (Table 8) showed no significant difference in mean response time for the four stimulus types between high-conflict (Mhigh-conflict = 646.79, SD = 10.21) and low-conflict conditions (Mlow-conflict = 634.94, SD = 9.60, p = 0.191). Baseline (Mbaseline = 658.10, SD = 9.17) and low-conflict conditions (Mlow-conflict = 634.94, SD = 9.60) differed significantly (p = 0.003). High-conflict (Mhigh-conflict = 646.79, SD = 10.21) and baseline conditions (Mbaseline = 658.10, SD = 9.17) did not differ significantly (p = 0.159).

PDP-C estimates for the gender Flanker task were analyzed via a 3 (previous task conflict level: high-conflict, baseline, low-conflict) × 2 (trait word category: masculine trait word, feminine trait word) repeated measures ANOVA. Results showed a significant main effect of previous task conflict level, $F(2, 56) = 6.58$, $p = 0.003$, $p^2 = 0.19$, but no significant main effect of trait word category or interaction ($ps > 0.05$). Thus, control processing estimates for the gender Flanker task across conflict levels did not differ between masculine and feminine trait word types. Further multiple comparisons indicated that the control processing estimate was significantly higher in the high-conflict condition ($M = 0.80$, $SD = 0.02$) than in the low-conflict condition ($M = 0.66$, $SD = 0.04$, $p = 0.001$)

and baseline condition (M = 0.70, SD = 0.03, $p = 0.009$). However, baseline and low-conflict conditions did not differ significantly ($p = 0.061$; see Table 8). High-conflict levels elicited the highest degree of control processing, low-conflict levels elicited the least, and baseline estimates fell in between. These results align with findings from Experiment 2 and the classification task preceding the Flanker task in this experiment. PDP-C essentially represents mean accuracy as a function of conflict level and prime stimulus. This pattern further supports our hypothesis that tasks under high-conflict levels can also elicit proactive control in subsequent tasks.

Finally, we analyzed differences in PDP-A estimates across conflict level tasks to test our predictions. Thirty-eight participants had valid PDP-A values across all gender Flanker task trial conditions and were included in this analysis. A 3 (conflict level: high-conflict, baseline, low-conflict) × 2 (prime stimulus: male picture, female picture) repeated measures ANOVA on PDP-A estimates revealed no significant main effects or interactions ($ps > 0.05$). Automatic processing estimates across conflict levels are presented in Table 8 , consistent with results from the preceding task.

Discussion

The primary purpose of Experiment 3 was to demonstrate whether cognitive control induced by different conflict levels exhibits cross-task consistency in its effect on gender stereotype expression. Therefore, ensuring differential cognitive control across conflict levels in the preceding task during task switching represents a critical prerequisite for testing this hypothesis (Egner, 2008; Freitas et al., 2007). Using a trial-to-trial control adjustment paradigm, we found that in the preceding word classification task, high-conflict conditions elicited proactive control and higher accuracy compared to baseline, thereby replicating Experiment 2 results. Further analysis of how classification tasks at different conflict levels influenced subsequent gender Flanker task processing revealed that proactive control activated in high-conflict classification tasks could induce cognitive control adjustments in subsequent tasks. Specifically, following high-conflict classification tasks, participants showed higher accuracy in the gender Flanker task. Moreover, control processing estimates calculated through PDP were higher for high-conflict levels than for low-conflict and baseline levels.

The proactive control processing advantage elicited by high-conflict conditions in the word classification task from Experiment 2 was validated once again, providing a foundation for cross-task consistency (Schmidt, 2019). This offers strong support for the notion that implicit stereotype tasks may alter conflict monitoring and cognitive control adjustments (Amodio et al., 2008; Freitas et al., 2007).

General Discussion

Task contexts with different conflict levels and cognitive control induced under different conditions produce varying effects on gender stereotype expression. Experiment 1 demonstrated that picture classification tasks suppressed gender stereotype expression regardless of conflict trial proportion. In contrast, Experiment 2’ s word classification tasks showed that high-conflict levels inhibited gender stereotype expression while low-conflict levels activated it, compared to baseline. Experiment 3 further revealed that proactive control generated under high-conflict levels could be maintained across subsequent tasks, unaffected by task type changes. These findings suggest that larger conflict trial proportions may be more effective for controlling stereotype expression, that control effectiveness relates to specific stereotype presentation formats, that individuals can sensitively detect subtle changes in conflict magnitude, and that conflict adaptation triggered by high-conflict tasks provides a control preparatory state for subsequent tasks (Bukowski et al., 2019; Schmidt, 2019).

5.1 Influence of Task Conflict Proportion on Gender Stereotype Expression

Classic cognitive conflict studies without social significance have demonstrated that experimental conditions with high proportions of conflict trials can induce a proactive control mode (Gratton et al., 1992). PDP analyses in Experiments 2 and 3 also found larger control processing values (PDP-C) under high-conflict versus low-conflict conditions (with baseline control estimates at intermediate levels). This indirectly demonstrates that high-conflict levels elicit proactive control rather than greater reactive activation of gender stereotypes. Counterintuitively, the finding that stronger interference environments lead to improved judgment accuracy aligns directly with proactive control mechanisms and previous research (Appelbaum et al., 2014). Further processing of automatic processing values (PDP-A) revealed that proactive and reactive control can cooperatively influence behavioral expression of implicit stereotypes rather than stereotype sources (e.g., automatically activated stereotypes). Other researchers have found that proactive and reactive control are not independent; when perceiving changes in conflict, individuals can simultaneously employ both cognitive control strategies to generate correct responses (Bugg & Braver, 2016). Our results fully support the notion that proactive and reactive control can complement each other according to situational demands.

5.2 Limitations on Stereotype Inhibition Under Task Conflict Proportions

Experiment 1’s results showed that conflict tasks with different trial proportions did not produce obvious differences in individuals’ cognitive processing. This outcome may relate to external presentation dimensions and internal processing formats of experimental materials. Experiment 2 replaced pictures with words as target stimuli, revealing that when gender pictures were presented under high-conflict conditions, proactive control elicited by abstract text-type target stimuli exerted some inhibitory effect on gender role stereotype expression.

First, both prime and target stimuli in Experiment 1 were colorful high-resolution pictures with high ecological validity. However, this picture format also possesses characteristics of being easily identifiable, memorable, and processable (Yuan et al., 2019). Theoretical perspectives on differences between text and image processing suggest that images tend toward bottom-up, concrete exemplar processing, whereas text processing involves more abstract processing than picture processing (Yuan et al., 2019). Images reduce top-down transformation resources at the cognitive level, enabling direct judgments based on concrete contexts (Sherman, 1996). Consequently, different conflict levels may have no impact on experimental results—that is, processing speed and accuracy for pictures may not differ across conflict levels. Experiment 2 validated the hypothesis that strong interference backgrounds elicit more proactive control than weak interference conditions, thereby reducing implicit stereotype expression (Amodio & Swencionis, 2018). Taken together, Experiments 1 and 2 indicate that proactive and reactive control processing elicited by different target stimulus types under high-interference environments are not entirely consistent, providing strong evidence that sensitivity to information presentation formats differs across conflict levels.

Building on Experiment 2' s conclusions, Experiment 3 used an improved trial-to-trial control adjustment paradigm to demonstrate that previous high-conflict-ratio gender stereotype classification tasks and subsequent gender Flanker tasks share similar cognitive control modes. Recent research has also demonstrated that conflict states can provide control preparation for subsequent tasks, thereby triggering stronger conflict adaptation in dynamic cognitive control processing (Kleiman et al., 2014; Bukowski et al., 2019; 张孟可 et al., 2021). Bukowski et al. (2019) used task-switching procedures involving social categories to assess cognitive flexibility. In their Experiment 1, participants categorized by gender or age; in Experiments 2 and 3, they categorized by gender or social role. They similarly found that participants' cognitive flexibility primarily manifested in task situations eliciting greater cognitive conflict, with cognitive flexibility across different task conflict proportions reflected in flexible adjustments of control strategies. Therefore, in gender Flanker tasks under high-conflict backgrounds, participants more readily elicited proactive control to suppress interference from Flanker stimuli, focusing attention on centrally located target stimuli. Under low-conflict backgrounds, they relied more on reactive control, extracting information from entire stimuli to generate responses.

Overall, extensive previous research has examined transitions in cognitive control flexibility but has been limited to comparisons within identical conflict task backgrounds, often neglecting control preparatory states between special presentation formats of information representation and different types of social stimuli. Since our daily behavioral activities occur in social conflict environments with varying proportions, considering task conflict proportion and social attributes is essential. In real life, traditionally normative gender role ideologies are not easily broken by new social norms, and group-based gender stereotypes and prejudices may be deeply rooted in people' s thinking and difficult to change. Under such circumstances, reducing attention to connections between gender and stereotypical information while focusing on target information in tasks may provide an effective method for weakening gender stereotype expression. Cognitive control triggered by different task conflict proportions can regulate implicit bias or stereotype expression to some degree. Moreover, conversion paradigms between different presentation formats of stimulus information and social meaning demonstrate that inhibitory control at the cognitive level can reduce bias or stereotype responses across specific presentation formats and tasks, making implicit bias or stereotype reduction and elimination not limited to reactive processing after stereotypes emerge.

5.3 Limitations and Future Directions

Despite important findings, this study has several limitations. First, cognitive control forms the basis of goal-directed behavior, yet how individuals adjust cognitive control for current goals based on different conflict task perceptions remains unknown, as does the distinction between proportion effects induced by practice effects versus those based on goals. Behavioral analysis alone cannot effectively answer these questions or ensure sufficient experimental results. Therefore, observing EEG activity during different conflict proportion tasks and identifying brain region differences in conflict monitoring and control adjustment would help us deeply investigate the cognitive neural mechanisms underlying conflict levels from an objective perspective, providing a theoretical foundation for effective cognitive control across different conflict environments. Second, conflict adaptation effects and proportion effects represent two manifestations of cognitive control. Conflict adaptation effects emphasize that individuals' monitoring of conflict in previous trials leads to better conflict control in current trials, whereas proportion effects focus more on conflict adaptation phenomena triggered by monitoring differences in conflict levels across trials within the same conflict task type. This study did not involve comparative analyses of these two effects nor deeply consider similarities and differences in cognitive control elicited by the two effects.

6 Conclusion

Cognitive control triggered by tasks with different conflict proportions can inhibit gender stereotype expression, though this effect is moderated by information presentation format. When perceived target information consists of image stimuli, task conflict ratio levels show no differential effect on gender stereotype expression. When perceived target information consists of semantic stimuli, high-conflict ratio tasks inhibit gender stereotype expression, whereas low-conflict ratio tasks enhance it. Moreover, cognitive control processing generated under high-conflict levels can be maintained across subsequent different tasks.

References

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartley, S. L., & Covert, A. E. (2004). Neural Signals for the Detection of Unintentional Race Bias. *Psychological Science, 15*(2), 88–93.

Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2008). Individual differences in the regulation of intergroup bias: The role of conflict monitoring and neural signals for control. *Journal of Personality and Social Psychology, 94*(1), 60–74.

Amodio, D. M., & Swencionis, J. K. (2018). Proactive control of implicit bias: a theoretical model and implications for behavior change. *Journal of Personality and Social Psychology, 115*(2), 255–275.

Appelbaum, L. G., Boehler, C. N., Davis, L. A., Won, R. J., & Woldorff, M. G. (2014). The Dynamics of Proactive and Reactive Cognitive Control Processes in the Human Brain. *Journal of Cognitive Neuroscience, 26*(5), 1021–1038.

Bailey, K., West, R., & Anderson, C. (2010). A negative association between video game experience and proactive cognitive control. *Psychophysiology, 47*(1), 34–42.

Bartholow, B. D., & Dickter, C. L. (2008). A response conflict account of the effects of stereotypes on racial categorization. *Social Cognition, 26*(3), 314–332.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. *Journal of Personality and Social Psychology, 70*(6), 1142–1163.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. J. Kane (Eds.) & A. Miyake & J. N. Towse (Ed.), *Variation in working memory* (pp. 76–106). Oxford University Press.

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106–113.

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 106*(18), 7351–7356.

Brown, J. W., Reynolds, J. R., & Braver, T. S. (2007). A computational model of fractionated conflict-control mechanisms in task-switching. *Cognitive Psychology, 55*(1), 37–85.

Bugg, J. M., & Braver, T. S. (2016). Proactive control of irrelevant task rules during cued task switching. *Psychological Research, 80*(5), 860–876.

Bukowski, M., de Lemus, S., Marzecová, A., Lupiáez, J., & Gocowska, M. A. (2019). Different faces of (un)controllability: Control restoration modulates the efficiency of task switching. *Motivation and Emotion, 43*(1), 12-34.

Chen, L. (2011). The schema and specificity of stereotype representation (Unpublished doctoral dissertation). Shanghai Normal University. [陈莉. (2011). 刻板印象的心理表征：组织结构及其特性 (博士学位论文). 上海师范大学.]

Croft, A. (2016). Men' s roles and women' s goals: causes, consequences, and complementarity (Unpublished doctoral dissertation). University of British Columbia.

Cui, Y. C. (2016). The cognitive control of impression formation on perception conflict. (Unpublished doctoral dissertation). Shanghai Normal University. [崔诣晨. (2016). 知觉冲突视阈下印象形成的认知控制 (博士学位论文). 上海师范大学.]

Cundiff, J. L., & Vescio, T. K. (2016). Gender stereotypes influence how people explain gender disparities in the workplace. *Sex Roles, 75*(3-4), 1-13.

Dickter, C. L., & Bartholow, B. D. (2010). Ingroup categorization and response conflict: interactive effects of target race, flanker compatibility, and infrequency on N2 amplitude. *Psychophysiology, 47*(3), 596-601.

Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, NJ: Erlbaum.

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.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences, 12*(10), 374-380.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses.* Behavior Research Methods, 41*, 1149-1160.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23*, 75-109.

Freitas, A. L., Bahar, M., Yang, S., & Banai, R. (2007). Contextual adjustments in cognitive control across tasks. *Psychological Science, 18*(2), 1040-1043.

Funes, M. J., Lupianez, J., & Humphreys, G. (2010b). Sustained vs. transient cognitive control: Evidence of a behavioral dissociation. *Cognition, 114*(3), 338-347.

Gonthier, C., Ambrosi, S., & Blaye, A. (2021). Learning-based before intentional cognitive control: developmental evidence for a dissociation between implicit and explicit control. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480-506.

Hazeltine, E., Lightman, E., Schwarb, H., & Schumacher, E. H. (2011). The boundaries of sequential modulations: Evidence for set-level control. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1898-1914.

Hilgard, J., Bartholow, B. D., Dickter, C. L., & Blanton, H. (2015). Characterizing switching and congruency effects in the Implicit Association Test as reactive and proactive cognitive control. *Social Cognitive and Affective Neuroscience, 10*(3), 381-388.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513-541.

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual Process Theories in Social Psychology* (pp. 383-400). New York, NY: Guilford Press.

Jia, L., Zhu, S. R., Zhang, C. J., & Zhang, Q. L. (2016). The distributed semantic representation and activation processes of the implicit and explicit stereotypes: An examination based on the perspective of cognitive neuroscience. *Advances in Psychological Science, 24*(10), 1519-1533. [贾磊, 祝书荣, 张常洁, 张庆林. (2016). 外显与内隐刻板印象的分布式表征及其激活过程——基于认知神经科学视角的探索. 心理科学进展, 24 (10), 1519-1533.]

Kidder, C. K., White, K. R., Hinojos, M. R., Sandoval, M., & Crites, S. L. (2018). Sequential stereotype priming: A meta-analysis. *Personality and Social Psychology Review, 22*(3), 199-227.

Kleiman, T., Hassin, R. R., & Trope, Y. (2014). The control-freak mind: Stereotypical biases are eliminated following conflict-activated cognitive control. *Journal of Experimental Psychology: General, 143*(2), 1-13.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., & Simon, S. (2016). Reducing implicit racial preferences: Intervention effectiveness across time. *Journal of Experimental Psychology: General, 145*(8), 1001-1016.

Liu, H. D. (2018). A research on gender role attitude, mate preference and the relationship between the two of college students (Unpublished master dissertation). Sichuan Normal University. [刘红豆. (2018). 大学生的性别角色态度与择偶偏好及其关系研究 (硕士学位论文). 四川师范大学.]

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167-202.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*(2), 181-192.

Purmann, S., Badde, S., Luna-Rodriguez, A., & Wendt, M. (2011). Adaptation to frequent conflict in the Eriksen flanker task. *Journal of Psychophysiology, 25*(2), 50-59.

Qiu, L., Zheng, X., & Wang, Y. F. (2008). Revision of the positive affect and negative affect scale. *Chinese Journal of Applied Psychology, 14*(3), 249-254. [邱林, 郑雪, 王雁飞. (2008). 积极情感消极情感量表 (PANAS) 的修订. 应用心理学, (03), 249-254.]

Rajadhyaksha, U., Korabik, K., & Aycan, Z. (2015). Gender, Gender-Role Ideology, and the Work-Family Interface: A Cross-Cultural analysis. In M. J. Mills (Ed.), *Gender and the Work-Family Experience: An Intersection of Two Domains* (pp. 1-15). Springer International Publishing.

Ren, N., Zuo, B., Hou, F. X., & Wang, G. J. (2012). Context effect or automatic process? The implicit attitude of undergraduates to old people. *Acta Psychologica Sinica, 44*(6), 777-788. [任娜, 佐斌, 侯飞翔, 汪国驹. (2012). 情境效应或自动化加工: 大学生对老年人的内隐态度. 心理学报, 44(06), 777-788.]

Rudman, L. A., & Kilianski. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin, 26*(11), 1315-1328.

Schmid, P. C., Kleiman, T., & Amodio, D. M. (2015). Neural mechanisms of proactive and reactive cognitive control in social anxiety. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 70*, 1-13.

Schmidt, J. R. (2019). Evidence against conflict monitoring and adaptation: An updated review. *Psychonomic Bulletin & Review, 26*, 753-771.

Sherman, J. W. (1996). Development and mental representation of stereotypes. *Journal of Personality & Social Psychology, 70*(6), 1126-1141.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*(2), 314-335.

Van, R. D., Van O. F., Vanhoomissen, T., Labiouse, C., & French, R. (2003). A recurrent connectionist model of group biases. *Psychological Review, 110*(3), 536-563.

Veen, V. V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI Study. *NeuroImage, 27*(3), 497-504.

Wang, P., Yang, Y. P., & Zhao, L. (2010). The activation of stereotypes: Behavioral and ERPs evidence. *Acta Psychologica Sinica, 42*(5), 607-617. [王沛, 杨亚平, 赵仑. (2010). 刻板印象的激活效应: 行为和 ERPs 证据. 心理学报, 42(5), 607-617.]

Wu, M. L. (2017). The relationship between egalitarian sex role attitudes of China dual-earner couples and their marital quality. (Unpublished master dissertation). Southwest University, Chongqing. [吴梦玲. (2017). 中国双薪夫妇的平等性别角色态度与婚姻质量的关系 (硕士学位论文). 西南大学, 重庆.]

Yuan, J., Tian, Y., Huang, X., Fan, H., & Wei, X. (2019). Emotional bias varies with stimulus type, arousal and task setting: Meta-analytic evidences. *Neuroscience & Biobehavioral Reviews, 107*, 461–472.

Yuan, F. (2015). Cognitive monitoring in different contexts: The effect of conflict adaptation on gender stereotype. (Unpublished master dissertation). Shaanxi Normal University. [袁菲. (2015). 认知监控的领域一般性：冲突适应对性别刻板印象的影响. (硕士学位论文). 陕西师范大学.]

Zhang, M. K., Li, Q., Yi, S. H., & Chen, A. T. (2021). Changes in the level of conflict trigger conflict adaptation. *Acta Psychologica Sinica, 53*(2), 128–138. [张孟可, 李晴, 尹首航, 陈安涛. (2021). 冲突水平的变化诱发冲突适应. 心理学报, 53(2), 128–138.]

Influence of Cognitive Control Based on Different Conflict Levels on the Expression of Gender Stereotypes

CHEN Li, SHI Xiao-ke, LI Wei-na, HU Yan
(School of Psychology, Northwest Normal University, Lanzhou 730030, China)

Abstract

Cognitive control is the ability of individuals to flexibly adjust their thoughts and behaviors and deal with compatible and conflicting information when facing tasks. It ensures that our actions are performed smoothly according to the expected purpose. However, conflict information contains pure cognitive conflict and social conflict information. With the development of society, the division of labor of gender roles in the family has changed. Nevertheless, the traditional stereotype that men work outside and women work inside the house still exists. Furthermore, although many studies have examined the consistency effect in gender stereotype activation, the influence of cognitive control induced under different conditions on gender stereotype expression is not apparent across task contexts varying in conflict levels. Therefore, based on previous research, the classic dual cognitive control theory that explains cognitive control processing, and the conflict monitoring theory that explains conflict tasks, this study systematically discusses the behavior patterns of gender stereotype expression under different conflict test times.

In Experiment 1, the participants were asked to complete the picture classification tasks with three conflict levels using the gender picture as the priming stimulus and housework picture and work picture as the target stimulus. Experiment 2 adjusted the target stimulus to semantic stimulus and asked the

participants to complete three-word classification tasks with different conflict levels. To further investigate the influence of varying conflict backgrounds of subsequent task conflict, Experiment 3 adopted the trial-to-trial control adjustment paradigm. In this paradigm, a full trial consists of two judgment tasks. First, the participants complete the word classification tasks with different conflict levels. Thereafter, they complete the gender Flanker tasks with the same conflict level.

Experiment 1 showed that the image classification task could inhibit the expression of gender stereotypes regardless of the proportion of conflict times. Regarding the word classification task in Experiment 2, compared to the baseline level, the expression of gender stereotypes was inhibited at the high conflict level and activated at the low conflict level.

Experiment 3 showed that the cognitive control processing generated by high conflict levels could be maintained in subsequent tasks. Furthermore, it was not affected by the change of task type. Additionally, Experiments 2 and 3 using the processing separation program showed that the control processing value under the high conflict condition was higher than that under the low conflict condition.

These results indicate that cognitive control induced by tasks with different conflict proportions can inhibit the expression of gender stereotypes; however, it is affected by the presentation of information.

Key words: gender stereotype, conflict processing, cognitive control strategy, proactive control, reactive control

Note: Figure translations are in progress. See original paper for figures.