Supplementary Material

Effects of stimulus type

A number of previous studies have investigated the effects of race on categorization speed. Typically, while individuals are better able recognize and identify same-race faces relative to other-race faces, they are able to classify other-race faces more quickly than same-race faces (e.g., Ge et al., 2009; Lewin, 1996; Zhao & Bentin, 2008). Additionally, numerous studies show increase P2 amplitude for other-race faces relative to same-race faces (for a review, see Bartholow & Dickter, 2011). Thus, we present mean differences in RT and P2 amplitude according to the type of face (Black/White, male/female, when applicable). Additionally, we examine the effect of fixation in Studies 2 and 3 and its effects of RT and P2 amplitude. For preliminary models testing effects of stimulus type on P2 amplitude and RT separately, we used the most complex random structure that was supported by the data (i.e., removing terms to avoid singular fit and terms that lead to high correlations between random effects; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017).

Study 1

RT. To test the effects of target race and fixation on response times, RT data were submitted to a multilevel model with target race (Black = −1, White = 1) and fixation (Eyes = −1, Forehead = 1) included as effect-coded predictors, along with the Target race × Fixation interaction.[1] The simple effects of both target race, $b = 2.98$, $t(59.6) = 2.43$, $p = .018$, and fixation, $b = 3.00$, $t(59.8) = 3.81$, $p < .001$, were significant. RTs were faster when categorizing Black faces compared to White faces and when fixating between the eyes compared to the

[1] Wilkinson notation: `RT ~ TarRace * Fix + (TarRace + Fix | Subject)`

forehead. The Target race x Fixation interaction was not significant, $b = -0.14$, $z = -0.20$, $p = .845$.

P2 amplitude. As reported in Volpert-Esmond, Merkle and Bartholow (2017), we tested the effects of race, fixation, and their interaction (effect-coded as before) on P2 amplitude.[2] The simple effect of race was significant, $b = -0.50$, $t(61.0) = -6.51$, $p < .001$, such that Black faces elicited larger P2 amplitudes than White faces. The simple effect of fixation was also significant, $b = -0.26$, $t(60.4) = -2.99$, $p = .004$, such that larger P2s were elicited in the eyes-fixation condition than the forehead-fixation condition. The Race x Fixation interaction was not significant.

Study 2

RT. Visual inspection of the means suggested the presence of a Target race x Target gender interaction in both tasks, such that RTs were faster to Black *male* faces than Black *female* faces, but that RTs were faster or equal to White *female* faces than White *male* faces. First, we confirmed the presence of this interaction in each task separately. Next, we tested its moderation by fixation condition and participant race. All models used effect coded variables: target race (Black = −1, White = 1), target gender (Female = -1, Male = 1), fixation (Eyes = −1, Forehead = 1), and participant race (Black = -1, White = 1).

Race-categorization task. The Target race x Target gender interaction was significant, $b = 7.12$, $z = 35.6$, $p < .001$, such that RTs were faster to Black male faces compared to Black female faces but slower (or equivalent) to White male faces compared to White female faces. However, the size of this interaction was significantly moderated by both participant race and fixation. For White participants, a significant Target race x Target gender x Fixation interaction, $b = -1.25$, $z =$

---

[2] Wilkinson notation: `P2amp ~ TarRace * Fix + (TarRace * Fix | Subject) + (1 | Electrode)`

-4.34, $p < .001$, indicated that the Target race x Target gender interaction was larger in the eyes condition, $b = 6.88$, $z = 17.1$, $p < .001$, than in the forehead condition, $b = 4.37$, $z = 10.7$, $p < .001$. For Black participants, the Target race x Target gender x Fixation interaction also emerged, $b = 1.12$, $z = 4.06$, $p < .001$, but patterns of means were in the opposite direction: the Target race x Target gender interaction was larger in the forehead condition, $b = 9.80$, $z = 25.1$, $p < .001$, than in the eyes condition, $b = 7.49$, $z = 19.3$, $p < .001$. The simple effect of fixation was not significant, $b = .69$, $t(60.58) = 0.89$, $p = .376$.

Gender-categorization task. As in the analysis of race-categorization task RTs, the Target race x Target gender interaction was significant, $b = 6.21$, $z = 27.6$, $p < .001$, and was significantly qualified by participant race and fixation. For White participants, the Target race x Target gender interaction was significant, $b = 3.80$, $z = 12.0$, $p < .001$, reflecting the same crossover pattern as in the race-categorization task. This was not qualified by an interaction with fixation. In contrast, for Black participants, the Target race x Target gender interaction was significantly moderated by fixation, $b = 1.29$, $z = 4.08$, $p < .001$. Breaking down this interaction revealed that the Target race x Target gender interaction was larger in the forehead condition, $b = 10.0$, $z = 22.6$, $p < .001$, than in the eyes condition, $b = 7.36$, $z = 16.40$, $p < .001$, similar to the pattern in the race-categorization task. The simple effect of fixation was not significant, $b = .44$, $t(60.28) = 0.66$, $p = .514$.

P2 amplitude. We included Target race (Black = −1, White = 1), target gender (female = −1, male = 1), fixation (eyes = −1, forehead = 1) and participant race (Black = −1, White = 1) as effect-coded predictors in separate models for each task.

Race-categorization task. As in the RT analyses, the Target race x Target gender interaction was significant, $b = -0.08$, $z = -3.77$, $p < .001$. This effect was not moderated by participant race, but was significantly moderated by fixation, $b = 0.15$, $z = 6.90$, $p < .001$. In the eyes condition, the Target race x Target gender interaction was significant, $b = -0.23$, $z = -7.64$, $p < .001$, such that P2s were significantly larger to Black *male* faces compared to Black *female* faces but significantly larger to White *female* faces compared to White *male* faces. In the forehead fixation condition, the Target race x Target gender interaction was also significant, $b = 0.06$, $z = 2.03$, $p < .001$; however, the pattern for the interaction differed from that in the eyes condition. P2s were larger to White male faces than White female faces, but did not significantly differ for Black male and female faces. Additionally, the simple effect of fixation was significant, $b = -0.32$, $t(60.9) = -4.02$, $p <.001$, such that fixating between the eyes elicited a larger P2 than fixating on the forehead.

Gender-categorization task. As in the race-categorization task, there was a significant Target race x Target gender x Fixation interaction, $b = -0.08$, $z = -3.56$, $p < .001$, but the Target race x Target gender x Participant race interaction was not significant, $b = 0.02$, $z = 0.90$, $p =.370$. In the eyes condition, the Target race x Target gender interaction was significant, $b = 0.11$, $z = 3.36$, $p =.001$, such that P2s were significantly larger to White male faces than White female faces, but there was no significant effect of gender for Black faces. In the forehead condition, the Target race x Target gender interaction was not significant, $b = -0.04$, $z = -1.41$, $p =.160$, nor was the effect of Target race, $b = -0.14$, $t(61.7) = -1.62$, $p =.110$, or Target gender, $b = 0.18$, $t(60.4) = 1.79$, $p =.079$. Additionally, the simple effect of fixation was significant, $b = -0.31$, $t(60.4) = -3.30$, $p =.002$, such that fixating between the eyes elicited a larger P2 than fixating on the forehead.

Study 3

RT. To test the effects of stimulus condition on reaction times, RT data were submitted to a multilevel model with target race (Black = −1, White = 1) and target gender (Female = -1, Male = 1) included as effect-coded predictors, along with the Target race x Target gender interaction.[3] This model was used to test RTs in the race-categorization and gender-categorization tasks separately.

Race-categorization task. The Target race x Target gender interaction was significant, b = 5.44, z = 17.1, p < .001. RTs to categorize Black male faces were faster than RTs to categorize Black female faces, b = -9.92, t(62.4) = -6.46, p < .001, but RTs to categorize White male and female faces did not differ significantly, b = 1.38, t(63.6) = 0.86, p = .394.

Gender-categorization task. As in the analysis of race-categorization task RTs, the Target race x Target gender interaction was significant, b = 2.92, z = 2.63, p =.011. However, the pattern of the interaction was different from that observed in the race-categorization task, such that RTs to categorize Black male and female faces did not differ significantly, b = -1.22, t(63.0) = -0.74, p = .464, whereas RTs to categorize White male faces were significantly slower than RTs to categorize White female faces, b = 4.63, t(63.0) = 2.65, p = .010.

P2 amplitude. We included Target race (Black = −1, White = 1) and target gender (female = −1, male = 1) as effect-coded predictors in separate models for each task.

Race-categorization task. Only the simple effect of target race was significant, b = -0.23, t(62.1) = -2.91, p = .005, such that P2s were larger to Black faces than White faces. The effects of target gender and the Target race x Target gender interaction were not significant.

[3] Wilkinson notation: `RT ~ TarRace * TarGender (TarRace * TarGender | Subject)`

Gender-categorization task. The simple effect of target race was marginally significant, $b = -0.19$, $t(62.1) = -2.00$, $p = .050$, such that P2s were larger to Black faces than White faces. The effects of target gender and the Target race x Target gender interaction were not significant.