Seeing is Not Stereotyping: The Functional Independence of Categorization and Stereotype Activation

Supplemental Materials

Main Study

Accuracy of Responses to Primes. At the end of each trial, after responding to the imperative stimulus, participants were probed to answer a block-appropriate question about the face prime to verify they processed the face primes differently by block. A 2 (Task: race categorization, feature detection) x 2 (Dot Presence: yes, no) x 2 (Prime Race: Black, White) x 2 (Target: gun, insect) within subjects ANOVA was performed on these responses. This revealed main effects of Task ($F(1,64)=7.51, p=.008, \eta_p^2=.105$) and Target ($F(1,64)=9.80, p=.003, \eta_p^2=.133$). Accuracy was greater when verifying the race of the prime (race categorization task) ($M=90.08\%, SD=.15$) compared to whether a dot was present (feature detection task) ($M=84.64\%, SD=.17$) and on trials in which the imperative stimulus was a gun ($M=87.80\%, SD=.14$) as compared to an insect ($M=86.92\%, SD=.14$), respectfully. There was also a Prime Race x Target interaction, $F(1,64)=5.15, p=.027, \eta_p^2=.075$. Simple effects tests showed that when primes were Black, participants were more accurate to answer questions about the primes when the imperative stimulus was a gun ($M=87.98\%, SD=.14$) as compared to an insect ($M=86.59\%, SD=.14$), $F(1,64)=16.25, p<.0001, \eta_p^2=.203$. There was no difference in accuracy as a function of imperative stimulus when responding to questions about White primes ($M_{guns}=87.63\%, SD=.14, M_{insects}= 87.25\%, SD=.14, F(1,64)=0.99, p=.321, \eta_p^2=.015$. There was also a marginally significant Task x Prime Race x Target interaction, indicating this latter effect was moderated by task, $F(1,64)=3.38, p=.070, \eta_p^2=.050$. Tests of the simple effect of Target within each level of Prime Race showed significant differences in accuracy to the prime on gun versus insect trials only for Black primes when the race categorization task was performed
(M_{g} = 91.02\%, SD = .14, M_{insects} = 89.00\%, SD = .15, F(1, 64) = 15.70, p < .0001, \eta^2_p = .197). None of the other simple effects of Target were significant (Fs = 0.05-2.35, p = .130-.813).

The Task main effect suggests the race categorization task may have been easier to perform than the feature detection task or, at least, the race of the prime was easier to remember than whether a dot was present at the time when questions were asked about the primes. The Target main effect and interaction with Prime Race are difficult to interpret. One possible mechanism is that identifying guns is easier than identifying insects, which in turn facilitated accuracy in recalling the status of the primes on gun trials. Such a possibility is consistent with repeated findings of greater accuracy to identify guns as compared to control objects in many implicit bias tasks (Correll, Park, Judd, & Wittenbrink, 2002; Judd, Blair, & Chapleau, 2004; Payne, 2001). Note, however, that we found no such Target main effect in accuracy in responses to the imperative stimulus in the present study. Another possibility suggested by the marginal Task x Prime Race x Target interaction is that stereotypes activated by the objects may have biased memory for the primes. In this case, seeing a gun as the imperative stimulus may have facilitated recall of the prime as Black in the race categorization task, selectively increasing accuracy in the race categorization task on gun trials when the prime was in fact Black.

We are unable to determine the precise mechanism behind the effects obtained in responses about the primes, but we emphasize the overall high accuracy levels in all conditions. This allows us to be confident that participants were in fact changing the way they processed the primes across blocks in accord with task instructions.

**Supplemental Study**

Supplemental data were collected to address the longer and less accurate responses to the imperative stimuli in the feature detection than racial categorization task in the main study. Full details on these supplemental data are given here.
Participants. Sixty University of Colorado undergraduates participated in exchange for course credit. The majority self-identified as White (n=55), with the remaining identifying as Asian (n=2), Hispanic (n=1), Persian (n=1), and multi-racial (n=1).

Procedure. The procedure was identical to the main study with two exceptions. First, the SOA between the prime and imperative stimulus was increased to 400 ms. Second, task order was counterbalanced for all participants, with an equal number of participants performing each task first. The purpose of this supplemental study was to verify that the absence of implicit stereotyping in the feature detection task in the main study was not simply a function of there being more participants who completed the feature detection task first in the main study, or that the feature detection task was more difficult. As such, our interest was in behavioral results, so all participants completed only the behavioral task, with no ERPs measured.

Results

Response Latencies to Imperative Stimulus. Response latencies were analyzed with a 2 (Task: race categorization, feature detection) x 2 (Dot Presence: yes, no) x 2 (Prime Race: Black, White) x 2 (Target: gun, insect) x 2 (Task Order: race categorization first, feature detection first) ANOVA, with all factors except the last manipulated within subjects. As in the main study, analyses were performance on log transformed latencies from correct trials, after dropping trials on which participants’ responses exceeded +/- 3 standard deviations relative to their mean.

Similar to the main study, there was a Task x Prime Race x Target interaction, although it was only of marginal significance in this study, \( F(1, 58)=3.38, p=.071, \eta_p^2=.055 \). As in the main study, we examined separate Prime Race x Target ANOVAs within each condition to determine task effects on implicit stereotyping. In the race categorization task, the predicted Target x Prime Race interaction was significant, \( F(1, 58)=9.11, p=.004, \eta_p^2=.136 \). Simple effects test showed that responses to guns were faster following Black (\( M=909.62 \text{ ms}, \))
SD = 399.73) than White primes ($M = 930.80$ ms, $SD = 413.39$), $F(1, 58) = 6.97$, $p = .011$, $\eta^2_p = .107$.

There was no difference in speed of responding to insects following Black ($M = 869.58$ ms, $SD = 353.30$) and White primes ($M = 857.83$ ms, $SD = 338.45$) in the race categorization task, $F(1, 58) = 2.28$, $p = .137$, $\eta^2_p = .038$. By contrast, the Prime Race x Target interaction was not significant within the feature detection task, $F(1, 58) = 0.14$, $p = .711$, $\eta^2_p = .002$.

In addition to the predicted Task x Prime Race x Target interaction, the omnibus analysis revealed four additional significant effects. One was the lower order Prime Race x Target interaction ($F(1, 58) = 7.10$, $p = .010$, $\eta^2_p = .109$) that was subsumed by the higher order Task x Prime Race x Target interaction just discussed. There were also main effects of Target ($F(1, 58) = 5.66$, $p = .021$, $\eta^2_p = .089$) and Block Order ($F(1, 58) = 9.03$, $p = .004$, $\eta^2_p = .135$). These showed that responses were faster to insects ($M = 855.36$ ms, $SD = 301.43$) than to guns ($M = 905.95$ ms, $SD = 352.32$) and faster from participants who performed the feature detection task first ($M = 754.11$ ms, $SD = 168.30$) compared to those who did the race categorization task first ($M = 1007.19$ ms, $SD = 387.42$). Task and Task Order also interacted, $F(1, 58) = 16.62$, $p < .0001$, $\eta^2_p = .223$. Tests of the simple effects of task within each Task Order condition show a pattern in which responses were faster during the second task performed. That is, among those who did the feature detection task first, responses were faster overall in the race categorization ($M = 732.96$ ms, $SD = 188.58$) than feature detection task ($M = 775.27$ ms, $SD = 161.12$), $F(1, 58) = 7.04$, $p = .010$, $\eta^2_p = .108$. By contrast, among those who did the race categorization task first, responses were faster overall on the feature detection ($M = 963.43$ ms, $SD = 358.60$) than race categorization task ($M = 1050.96$ ms, $SD = 435.02$), $F(1, 58) = 9.69$, $p = .003$, $\eta^2_p = .143$.

**Accuracy to Imperative Stimulus.** With the longer SOA, responses to the target were quite accurate. Nevertheless, a 2 (Task: race categorization, feature detection) x 2 (Dot Presence: yes, no) x 2 (Prime Race: Black, White) x 2 (Target: gun, insect) x 2 (Task Order: race categorization first, feature detection first) ANOVA showed modulation of implicit
stereotyping by task. The expected three-way Task x Prime Race x Target interaction was significant, $F(1,58)=3.99, p=.050, \eta^2_{p}=.064$. Separate Prime Race x Target ANOVAs within each task condition revealed the expected significant interaction in the race categorization condition, $F(1,58)=6.39, p=.014, \eta^2_{p}=.098$. Participants tended to more accurately categorize guns following Black ($M=99.0\%, SD=.02$) than White primes ($M=98.6\%, SD=.03$), but this difference was only of marginal significance, $F(1,58)=3.76, p=.057, \eta^2_{p}=.060$. There were no differences in accuracy to insects as a function of prime race in the race categorization task ($F(1, 58)=1.62, p=.209, \eta^2_{p}=.027$), and no Prime Race x Target Type effect in the feature detection task ($F(1,58)=0.31, p=.580, \eta^2_{p}=.050$).

There were 3 additional significant effects in the omnibus ANOVA. The main effects of Task ($F(1,58)=4.07, p=.048, \eta^2_{p}=.066$) and Dot Presence ($F(1,58)=4.58, p=.037, \eta^2_{p}=.073$) indicate that participants were more accurate overall during the race categorization ($M=98.8\%, SD=0.01$) than feature detection block ($M = 98.0\%, SD = 0.04$) and when pictures lacked dots ($M=98.6\%, SD=0.02$) as compared to when they had dots on them ($M=98.3\%, SD=0.02$), respectively. There was also a Task x Task Order interaction, $F(1,58)=15.08, p<.0001, \eta^2_{p}=.206$. Tests of the simple effect of task within each Task Order condition revealed greater accuracy in the race categorization ($M=99.2\%, SD=0.01$) than feature detection task ($M=96.7\%, SD=0.01$) among those who did the race categorization task first, $F(1,58)=17.41, p<.001, \eta^2_{p}=.23$. For those who did the feature detection task first, there was no reliable difference in accuracy between the feature detection task ($M=99.3\%, SD=0.01$) and the race categorization ($M=98.5\%, SD=0.02$), $F(1,58)=1.74, p=.192, \eta^2_{p}=.029$. When considered with the response latency results, these results may indicate a speed-accuracy trade-off as the study progressed for participants who did the race categorization task first (they got faster but less accurate in the second task they performed).
Accuracy of Responses to Primes. Accuracy to the questions about the primes answered at the end of each trial were analyzed with a 2 (Task: race categorization, feature detection) x 2 (Dot Presence: yes, no) x 2 (Prime Race: Black, White) x 2 (Target: gun, insect) within subjects ANOVA. As with the main study, accuracy rates were generally high in all conditions, indicating that participants were attentive to race during the race categorization task and to the race-irrelevant dot during the feature detection task. As with the main study, there was a main effect of Task, with greater accuracy to verify prime race (race categorization task) \((M=93.77\%, \ SD=.05)\) than dot presence (feature detection task) \((M=88.21\%, \ SD=.12)\), \(F(1,59)=17.63, \ p<.0001, \ \eta_p^2=.230\). There was also a Dot Presence main effect showing greater accuracy for faces with dots \((M=91.48\%, \ SD=.07)\) than those without dots \((M=90.49\%, \ SD=.08)\), \(F(1,59)=8.46, \ p=.005, \ \eta_p^2=.125\). These main effects were moderated by the Task x Dot Presence interaction, \(F(1,59)=7.30, \ p=.009, \ \eta_p^2=.110\). The simple effect of task, with greater accuracy to verify the race of the prime than whether a dot was present, was significant both when the prime had a dot \((M_{\text{race task}}=93.73\%, \ SD=.06 \text{ and } M_{\text{dot task}}=89.23\%, \ SD=.11)\), \(F(1,59)=12.25, \ p=.001, \ \eta_p^2=.172\) and when it lacked a dot \((M_{\text{race task}}=93.80\%, \ SD=.05 \text{ and } M_{\text{dot task}}=87.19\%, \ SD=.13)\), \(F(1,59)=20.28, \ p<.0001, \ \eta_p^2=.256\). Instead, the interaction is attributable to a significant simple prime dot effect within the feature detection task, but not the race categorization task. That is, when performing the feature detection task, participants were more accurate to recall the status of the prime when it had a dot on it \((M=98.23\%, \ SD=.11)\), as compared to when it lacked a dot \((M=87.19\%, \ SD=.13)\), \(F(1,59)=10.24, \ p=.002, \ \eta_p^2=.148\). By contrast, when performing the race categorization task, there was no difference in accuracy when responding to primes with \((M=93.73\%, \ SD=.06)\) and without dots \((M=93.80\%, \ SD=.05)\), \(F(1,59)=0.04, \ p=.84, \ \eta_p^2=.001\).

The only other significant effect was the Prime Race x Target interaction, \(F(1,59)=7.73, \ p=.007, \ \eta_p^2=.116\). Unlike the main study, tests of simple effects showed that when the prime
was Black, participants were more accurate in their recall of it when the imperative stimulus was an insect (\(M=91.35\%, SD=.08\)) as compared to a gun (\(M=89.82\%, SD=.09\)), \(F(1,59)=5.11, p=.02, \eta_p^2=.088\). By contrast, accuracy was equally high to White primes when the imperative stimulus was a gun (\(M=91.36\%, SD=.08\)) and an insect (\(M=91.41\%, SD=.08\)), \(F(1,59)=0.008, p=.927, \eta_p^2<.0001\).

**Discussion**

The most critical finding from this supplemental study is that the pattern of implicit stereotyping observed in the main study – Black primes facilitating responses to guns in the race categorization but not feature detection block -- is reliable and replicable. When performing the race categorization task, participants were faster and (marginally) more accurate to identify guns following Black than White primes. Not such differences to guns emerged when participants were performing the feature detection task. Moreover, the supplemental study suggests that task effects in the main study were not an artifact of either a practice effect or large differences in task difficulty. To review, we found in the main study that participants were faster and more accurate to the imperative stimulus in the race categorization than feature detection task. One possible explanation of such a pattern is that the feature detection task is more difficult than the race categorization task: potentially scanning a great deal of the image in search of a dot that may or may not be present could be more difficult than identifying race. We do not find such a possibility overly problematic. The purpose of the feature detection task is to provide participants with a goal that does not require attention to face race in order to complete. It was not necessary that the task be equal to race detection in difficulty. Moreover, task difficulty is not a particularly parsimonious explanation of our effects in light of other results in the main study. In particular, if the feature detection task was so much more difficult, we might expect different ERP responses than the ones obtained. Most notably, if the feature detection was more difficult than the race task, that would likely interfere with participants’ ability to
encode race in the feature detection task (e.g., race effects might be completely absent or delayed in the P200 or N200). This was not the case. Instead, identical ERP race effects occurred in the feature detection task and race categorization task.

Another possible explanation for faster and more accurate responses to the imperative stimulus in the race categorization task in the main study is that participants’ performance improved over the course of the study. This is a possible explanation because while we varied task order between subjects in the main experiment among those who had ERPs recorded, a programming error led all participants who provided only behavioral data to perform the feature detection task first. Overall, then, there were more participants in the main study who did the feature detection task first. The better performance on the race categorization task might then reflect a practice effect. Again, we do not think this possibility is problematic for our substantive conclusions because if the feature detection task was more difficult because it was performed first by more participants, that might work to decrease sensitivity to race in the ERPs, yet we find racial modulation in the P200 and N200 in the feature detection task.

Nevertheless, to be conservative, we wanted to verify that the pattern of implicit stereotyping we observed in the main study is procedurally robust. That we obtained the critical task modulation of implicit stereotyping in this supplemental study confirms that the behavioral effects reported in the main study are not an artifact of task difficulty or practice effects. When we diminish the demand of scanning for a dot by increasing the SOA, and remove the unintended order confound by systematically varying task order, we obtain the same pattern of implicit stereotyping as the main study.

To the issue of what caused the slower and less accurate responses in the feature detection task in the main study, we suspect it may be a combination of dot detection taking more time and a practice effect. We say this because there is some evidence of a practice effect in the supplemental data in that response latencies decreased in the second task participants
performed. This was true in both task orders. However, this decrease in speed was not accompanied by an increase in accuracy. To the contrary, for those who did the race categorization task first, their decrease in speed in the feature detection (i.e., second) block was also accompanied by a decrease in accuracy. For those who did the feature detection task first, the decrease in reaction time in their second task (race categorization) was not accompanied by a decrease in accuracy. Thus, the evidence for task improvement over time is equivocal. At the same time, we slightly increased the SOA in this supplemental study. This would have allowed participants more time to complete their dot evaluation before needing to respond to the imperative stimulus, thereby allowing them to respond more quickly and accurately to the guns and insects.