

Running Head: Stereotypes and Response Inhibition

Stop Stereotyping

Johanna K. Falbén,¹ Juliana L. Olivier,² Marius Golubickis,³ Nerissa S. P. Ho,²
Linn M. Persson,¹ Dimitra Tsamadi,¹ Eleni Marinopoulou,¹ Bianca Bianciardi,¹
William A. Cunningham,³ C. Neil Macrae¹

¹School of Psychology, University of Aberdeen, Aberdeen, Scotland, UK

²Department of Psychology, University of York, York, England, UK

³Department of Psychology, University of Toronto, Toronto, Canada

Address Correspondence To:

Johanna Falbén
School of Psychology
University of Aberdeen
Aberdeen AB24 3FX
Scotland, UK

Email: r01jfk17@abdn.ac.uk

Abstract

Restraining the expression of stereotypes is a necessary requirement for harmonious living, yet surprisingly little is known about the efficacy of this process. Accordingly, in two experiments, here we used a stop-signal task to establish how effectively stereotype-related responses can be inhibited. In Experiment 1, following the presentation of gender-typed occupational contexts, participants reported the sex of target faces (i.e., Go trials) unless an occasional auditory tone indicated they should withhold their response (i.e., Stop trials). In Experiment 2, following the presentation of male and female faces, participants made either stereotypic or counter-stereotypic judgments, unless a stop signal was presented. Regardless of whether stereotyping was probed indirectly (Expt. 1) or directly (Expt. 2), a consistent pattern of results was observed; inhibition was faster for stereotypic compared to counter-stereotypic responses. These findings demonstrate that stopping stereotyping may be less challenging than has widely been assumed.

Keywords: stereotyping, response inhibition, person construal, executive control

Stop Stereotyping

With depressing regularity, individuals respond toward others in racist, sexist, homophobic, or otherwise disagreeable stereotype-related ways. These actions, moreover, occur in even the most visible of arenas. For example, in 2011, two prominent male sports commentators in the U.K. were fired for remarking that female officials are incapable of understanding the offside rule in soccer. Extending beyond the obvious offence that comments such as these provoke, the wider impacts of stereotyping (in its various forms) are insidious and tangible. For example, in the U.K. alone, the economic cost of prejudice in the workplace is estimated to be around £127 (~\$163) billion annually (Cebr, 2018). Worldwide, discriminatory practices are reckoned to reduce global income by around 16% (i.e., \$12 trillion, OECD, 2016). Given that the purveyors of stereotyping — be it an ageist remark in the office or a wolf whistle by the pool — are routinely humiliated and embarrassed by their actions, an interesting question arises. If the personal and societal costs of stereotyping are so great, why do people not just stop themselves from responding in this way?

Although there are multiple reasons why stereotyping persists (Fiske, 1998), one in particular dominates contemporary psychological theorizing; stereotype-based responding saves people the trouble of thinking deeply about others. Supported by decades of research, the application of stereotypes has been shown to economize (i.e., streamline) core aspects of social-cognitive functioning. Most notably, compared to counter-stereotypic material, stereotype-consistent stimuli are detected with rapidity, processed efficiently, and exert undue influence on judgment and memory (Bodenhausen & Macrae, 1998; Brewer, 1988; Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Hilton & von Hippel, 1996; Macrae & Bodenhausen, 2000). Serving as repositories of culturally shared knowledge (e.g., men are aggressive, librarians are introverted, Scots are stingy), stereotypes furnish person-related information without the cumbersome necessity of social interaction. In other words, as simplifying cognitive tools, category-related beliefs facilitate both information processing and response generation. Given therefore the benefits of

categorical thinking, it is perhaps understandable why stereotyping prevails. The ease with which stereotypic responses can be generated and executed may simply make them difficult to stop (Bargh, 1999; Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Macrae & Bodenhausen, 2000). As Bargh (1999, p. 378) has argued, “Once a stereotype is so entrenched that it becomes activated automatically, there is relatively little that can be done to control its influence.” But is this actually the case?

As things currently stand, the inhibition of stereotype-related responses remains poorly understood. To date, efforts to explicate the regulation of stereotyping have focused almost entirely on how readily stereotypic thoughts can be banished (i.e., suppressed) from consciousness (e.g., Macrae, Bodenhausen, Milne, & Jetten, 1994; Monteith, Sherman, & Devine, 1998). Less conspicuous in the literature is work exploring the intentional stopping of stereotype-based deeds (but see Bartholow, Dickter, & Sestir, 2006). This oversight is surprising as response inhibition (i.e., stopping an already initiated, but uncompleted, action) is a core component of executive function and has been studied extensively elsewhere (Diamond, 2013; Friedman & Miyake, 2004). Moreover, in laboratory settings, the stop-signal task has emerged as the primary method for exploring this ability (Verbruggen & Logan, 2008). In this task, participants are presented with a visual stimulus (i.e., Go signal) that requires a speeded motor response. Occasionally, after a variable delay (i.e., stop-signal delay, SSD), the Go stimulus is followed by a stop signal (e.g., an auditory tone) instructing participants that the response should be withheld. While, at short SSDs, response inhibition is usually successful; at longer delays inhibition often fails and the action is performed.¹ The utility of this paradigm is that it enables estimation of the covert latency of the stop process — the stop-signal response time (SSRT, see Logan & Cowan, 1984; Logan, Cowan, & Davis, 1984). That is, the efficiency of response inhibition can be established.

¹ Logan and Cowan (1984) proposed a horse-race model to account for results in this paradigm. The model assumes there are two independent processes (i.e., Go process & Stop process) with stochastically independent finishing times.

Using a standard stop-signal task, here we considered the ease with which stereotyped (i.e., stereotypic/counter-stereotypic) responses can intentionally be stopped.² Following the presentation of gender-typed occupational contexts, participants were required to report the sex of target faces (i.e., Go trials) unless an auditory tone signalled they should withhold their response (i.e., Stop trials). Sequential priming tasks such as these are commonly used to activate stereotype-related knowledge in memory (Kidder, White, Hinojos, Sandoval, & Crites, 2018; Wentura & Rothermund, 2014). Although it is tempting to presume that stereotypic responses must be difficult to stop (Bargh, 1999; Bartholow et al., 2006), there is good reason to suspect that the opposite may be the case — stereotype-consistent responses may be inhibited more effectively than stereotype-inconsistent responses. In response-priming paradigms of the type used here, prime and target stimuli elicit either compatible (i.e., consistent trials) or incompatible (i.e., inconsistent trials) responses (Wentura & Rothermund, 2014). Thus, on stereotype-inconsistent (vs. stereotype-consistent) trials, participants must not only potentially deal with a stop signal (i.e., suppress a motor response), but also with pre-activation of the incorrect (i.e., primed) response. As a result of these increased inhibitory demands, SSRTs should be elevated to counter-stereotypic compared to stereotypic stimuli (see Kramer, Humphrey, Larish, Logan, & Strayer, 1994; Ridderinkhof, Band, & Logan, 1999; Verbruggen, Liefoghe, & Vandierendonck, 2004, 2006). We explored this hypothesis in the current experiment.

² Although Bartholow et al. (2006) also used a stop-signal task to explore the inhibition of stereotype-related responses, SSRTs were not estimated in this research, instead emphasis centered on how alcohol triggers stop-signal failures (i.e., failed inhibition).

Experiment 1

Method

Participants and Design

Thirty undergraduates (10 males, $M_{\text{age}} = 22.27$, $SD = 2.82$) took part in the experiment.³ All participants had normal or corrected-to-normal visual acuity. Two participants (2 females) failed to follow the instructions, thus were excluded from the analysis. Informed consent was obtained from participants prior to the commencement of the experiment and the protocol was reviewed and approved by the Ethics Committee at the School of Psychology, University of Aberdeen. The experiment had a 2 (Context: auto-repair workshop or cosmetics store) X 2 (Target: male or female) repeated measures design.

Stimulus Material and Procedure

Participants arrived at the laboratory individually, were greeted by the experimenter, seated in front of a desktop computer, and told they would be performing a sex-categorization task. In the task, a gender-typed occupational context was presented (i.e., auto-repair workshop or cosmetics store), followed by either a male or a female face. Using two buttons on the keyboard (i.e., N & M), participants had to report the sex of each face (Falbén et al., in press). The faces were taken from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) and were 140 x 176 pixels in size, greyscale, and depicted young adults aged 20-30 years. The context pictures were taken from Google images, were 500 x 500 pixels in size, greyscale, and contained no people.

³ Based on a medium effect size, G*Power ($\eta_p^2 = .05$, $\alpha = .05$, power = 80%) revealed a requirement of 28 participants. An additional ~10% were recruited to allow for drop out.

Each trial began with the presentation of a central fixation cross for 1000 ms, followed by a pictorial context (i.e., auto-repair workshop or cosmetics store) which remained on screen for 500 ms, after which it disappeared and was replaced by a male or female face for 50 ms. The screen then turned blank and participants had to report, by pressing the appropriate button on the keyboard as quickly and accurately as possible, whether the target was male or female. Participants had 1500 ms to make a response (6% of responses fell outside the response window). Eighty faces (40 male & 40 female) were used, with each face preceded by both contexts. The meaning of the response buttons was counterbalanced across the sample. Participants initially performed 20 practice trials, followed by 4 blocks each containing 160 experimental trials in which context-consistent (i.e., auto-repair workshop/male face, cosmetics store/female face) and context-inconsistent (i.e., auto-repair workshop/female face, cosmetics store/male face) stimulus pairs appeared equally often in a random order. Critically, on 25% of the trials a stop signal (i.e., an auditory beep) was presented, indicating that participants should withhold their response (i.e., do not press the button). The stop signal was presented at 1000 Hz for 100 ms and occurred dynamically. Specifically, SSDs were increased or decreased by 50 ms when the participant succeeded or failed to inhibit, respectively. The SSD values were drawn from four interleaved staircases, resulting in 40 trials from each staircase for a total of 160 stop trials (i.e., 40 stop trials per block). Stimulus and response events were presented using Matlab (Mathworks) and the Psychtoolbox (www.psychtoolbox.org). On completion of the task, participants were debriefed, thanked, and dismissed.

Results

A 2 (Context: auto-repair workshop or cosmetics store) X 2 (Target: male or female) analysis of variance (ANOVA) was conducted on participants' mean Go response times (Go RTs), response accuracy, SSDs, failed inhibition, and SSRTs, the results of which are summarized below (see Table 1).

Go RTs. Responses faster than 200 ms were excluded from the analysis, eliminating less than 1% of the overall number of trials. The only effect to emerge in the analysis was a significant Context X Target interaction, $F(1, 27) = 51.05, p < .001, \eta_p^2 = .65$. Further analysis of the interaction revealed a significant consistency effect for both the auto-repair ($t(27) = 5.34, p < .001, d_z = 1.01$) and cosmetics ($t(27) = 6.11, p < .001, d_z = 1.15$) contexts, such that responses were faster on context-consistent (i.e., stereotypic) than context-inconsistent (i.e., counter-stereotypic) trials.

Accuracy. The analysis yielded no significant effects.

SSDs. The analysis revealed no significant effects.

Failed Inhibition. No significant effects were observed.

SSRTs. SSRTs were estimated using the quantile method, which does not rely on the assumption of a 50% inhibition failure rate (Band, van der Molen, & Logan, 2003). To calculate quantile SSRTs, all Go RTs were arranged in ascending order and then the Go RT corresponding to the observed inhibition failure rate was selected, yielding the quantile RT. The average SSD was subtracted from this quantile RT, providing an estimate of the SSRT. The only effect to emerge in the analysis was a significant Context X Target interaction, $F(1, 27) = 9.97, p = .004, \eta_p^2 = .27$. Further analysis of the interaction revealed a significant consistency effect for both the auto-repair ($t(27) = 3.02, p = .005, d_z = 0.57$) and cosmetics ($t(27) = 2.47, p = .020, d_z = 0.47$) contexts, such that SSRTs were shorter on context-consistent (i.e., stereotypic) than context-inconsistent (i.e., counter-stereotypic) trials.⁴

⁴ No association between Go RTs and SSRTs was observed, $r(27) = -.040, p = .835$.

Table 1. Task performance as a function of Context and Target (Experiment 1)

Measure	Auto-repair Workshop		Cosmetics Store	
	Male (consistent)	Female (inconsistent)	Male (inconsistent)	Female (consistent)
Go RT (ms)	734 (30)	776 (28)	767 (25)	710 (27)
Accuracy (%)	86 (2)	86 (2)	87 (2)	88 (2)
SSD (ms)	257 (15)	254 (15)	253 (15)	256 (14)
Failed inhibition (%)	36 (4)	32 (4)	37 (4)	30 (4)
SSRT (ms)	357 (22)	394 (30)	400 (41)	339 (28)

Note. RT = reaction time. Standard error of the mean (SEM) in parentheses.

Hierarchical Drift Diffusion Modeling

To identify the operations through which contextual expectancies impacted task performance, Go trials were submitted to an additional hierarchical drift diffusion model (HDDM) analysis (Ratcliff, Smith, Brown, & McKoon, 2016; White et al., 2014; Wiecki, Sofer, & Frank, 2013). Elucidating the processes that underpin performance on Go trials in a sequential priming paradigm provides potentially useful insight into the dynamics of response inhibition (see Supplemental Material for a description of drift diffusion modeling and the current analysis).

Interrogation of the posterior distributions for the best fitting model demonstrated that performance on Go trials was underpinned by a response bias. Specifically, contextual expectancies shifted the starting point (z) of evidence accumulation, such that the auto-repair context was closer to the upper threshold (i.e., male response, $z = .51$), and the cosmetics context was closer to the lower threshold (i.e., female response, $z = .48$). There was strong evidence that these starting values

differed from each other ($p_{\text{Bayes}}[\text{auto-repair} > \text{cosmetics}] = .018$).⁵ In addition, the auto-repair context yielded suggestive evidence for a bias in starting value (comparison with $z = .50$, $p_{\text{Bayes}}[\text{auto-repair} > .50] = .128$), and the cosmetics context strong evidence for such a bias ($p_{\text{Bayes}}[\text{cosmetics} < .50] = .019$). No evidence for a difference in non-decision processes (t_0) was observed ($p_{\text{Bayes}}[\text{context-consistent} < \text{context-inconsistent}] = .271$, respective M s: .374 s vs. .391 s). These results confirm that prior expectancies impacted person construal via a shift in the starting point of evidence accumulation (i.e., pre-activation of prime-consistent responses).

Discussion

The results of Experiment 1 support the hypothesis that it is easier to inhibit stereotypic compared to counter-stereotypic responses (cf. Bargh, 1999; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000). In a sequential-priming paradigm, SSRTs were faster when target stimuli (i.e., faces) were consistent rather than inconsistent with respect to the preceding occupational context. In addition, an HDDM analysis revealed that prior expectancies influenced task performance via a response bias (Ratcliff et al., 2016). The demonstration that stopping was faster for stereotypic than counter-stereotypic responses corroborates previous work demonstrating that inhibition is facilitated when targets are accompanied by response-congruent compared to response-incongruent stimuli (Kramer et al., 1994; Ridderinkhof et al., 1999; Verbruggen et al., 2004, 2006).

Given the counterintuitive flavor of the current results (i.e., it is easier to inhibit stereotypic rather than counter-stereotypic responses), the goal of our second experiment was quite straightforward; to attempt to replicate the effects observed in Experiment 1 using a different paradigm and measure of stereotyping. Compared with a sequential priming task in which

⁵ Bayesian p values quantify the degree to which the difference in the posterior distribution is consistent with the hypothesis that the parameter is greater for stereotypes than counter-stereotypes. For example, a Bayesian p of .05 indicates that 95% of the posterior distribution support the hypothesis.

stereotyping was assessed indirectly (i.e., Expt. 1 – participants provided sex-categorization judgments); in our next experiment, participants were required to make explicit stereotype-based responses (Osterhout, Bersick, & McLaughlin, 1997; Wang, Tan, Zhang, Wang, & Luo, 2018; White, Crites Jr., Taylor, & Corral, 2009). Specifically, following the presentation of a male or female face, participants had to provide either stereotypic or counter-stereotypic responses on the basis of occupational or trait-based stereotype-related information. Replicating Experiment 1, we expected response inhibition to be faster for stereotypic compared to counter-stereotypic responses.

Experiment 2

Method

Participants and Design

Eighty undergraduates (14 males, $M_{\text{age}} = 21.85$, $SD = 3.54$) took part in the experiment.⁶ All participants had normal or corrected-to-normal visual acuity. Two participants (2 females) failed to follow the instructions, thus were excluded from the analysis. Informed consent was obtained from participants prior to the commencement of the experiment and the protocol was reviewed and approved by the Ethics Committee at the School of Psychology, University of Aberdeen. The experiment had a 2 (Block Type: stereotypic or counter-stereotypic) X 2 (Target: male or female) X 2 (Judgment: occupation or trait) mixed design with repeated measures on the first and second factors.

⁶ Based on a medium effect size, G*Power revealed a requirement of 80 participants.

Stimulus Materials and Procedure

Participants arrived at the laboratory individually, were greeted by the experimenter, seated in front of a desktop computer, and told they would be performing a social-judgment task. First, participants were randomly assigned to make judgments pertaining to either occupational (i.e., mechanic vs. hairdresser) or trait-related (i.e., dominant vs. caring) stereotype-based information. This manipulation served as a between-participants replication of the effect of interest. Next, participants were told that male or female faces would appear on the computer screen and their task was simply to respond to them in a stereotypic or counter-stereotypic manner on the basis of prevailing occupational or personality-related stereotypes. For the stereotypic block of trials, participants were instructed to respond to all male faces with the occupation ‘mechanic’ (i.e., occupation condition) or the trait ‘dominant’ (i.e., trait condition) and, correspondingly, to all female faces with the occupation ‘hairdresser’ or the trait ‘caring.’ On the counter-stereotypic block of trials, this face-response mapping was reversed (i.e., respond to male faces with ‘hairdresser’ or ‘caring’ and to female faces with ‘mechanic’ or ‘dominant’). Critically, on 25% of the trials, a stop signal (i.e., auditory beep) was presented, indicating that participants should withhold their response.

Each trial began with the presentation of a central fixation cross for 500 ms, followed by a male or female face for 1000 ms. Participants had to make either occupational or trait-related stereotype-based judgments as quickly and accurately as possible using two buttons on the keyboard (i.e., N & M). A response was required while each face remained on the screen (3% of responses fell outside the response window). Sixty faces (30 male & 30 female) were taken from the Chicago Face Database (Ma et al., 2015). Participants initially performed 20 practice trials, followed by four blocks of 120 experimental trials in which two blocks were stereotypic (i.e., males are mechanics/dominant; females are hairdressers/caring) and two blocks were counter-stereotypic (i.e., males are hairdressers/caring; females are mechanics/dominant). Block order and the meaning

of the response buttons was counterbalanced across the sample. The stop-signal procedure was as in Experiment 1. On completion of the task, participants were debriefed, thanked, and dismissed.

Results

A 2 (Block Type: stereotypic or counter-stereotypic) X 2 (Judgment: occupation or trait) mixed model ANOVA was conducted on participants' mean Go RTs, response accuracy, SSDs, failed inhibition, and SSRTs, the results of which are summarized below (see Table 2).⁷

Go RTs. Responses faster than 200 ms were excluded from the analysis, eliminating less than 1% of the overall number of trials. The only effect to emerge in the analysis was a main effect of Judgment [$F(1, 77) = 7.99, p = .006, \eta_p^2 = .09$], such that RTs were faster when responses were based on occupational ($M = 568$ ms, $SD = 67$ ms) compared to trait-related information ($M = 611$ ms, $SD = 74$ ms).⁸

Accuracy. The analysis yielded a main effect of Block Type [$F(1, 77) = 11.55, p < .001, \eta_p^2 = .13$], indicating that responses were more accurate in the stereotypic ($M = 96\%$, $SD = 4\%$) rather than counter-stereotypic ($M = 94\%$, $SD = 6\%$) block.

SSDs. The analysis revealed a main effect of Block Type [$F(1, 77) = 5.42, p = .023, \eta_p^2 = .07$], revealing that SSDs were longer in the stereotypic ($M = 269$ ms, $SD = 54$ ms) compared to the counter-stereotypic ($M = 261$ ms, $SD = 54$ ms) block.

Failed Inhibition. The analysis yielded a main effect of Block Type [$F(1, 77) = 7.15, p = .009, \eta_p^2 = .08$], indicating that participants failed to withhold their responses less often when a stop

⁷ Preliminary analysis revealed no effects of target, consequently data were collapsed across this factor.

⁸ Unlike Experiment 1, Go RTs were not speeded for stereotypic compared to counter-stereotypic responses. However, it is commonplace in stop-signal tasks for participants to engage in strategic slowing (i.e., delaying responses) during Go trials, as this increases the probability of successful inhibition (Bissett & Logan, 2011; Verbruggen & Logan, 2009). That slowing on Go trials only occurred in Experiment 2 may reflect the focus on explicit stereotyping in this study.

signal was presented during the stereotypic ($M = 31\%$, $SD = 14\%$) rather than counter-stereotypic ($M = 34\%$, $SD = 14\%$) block.

SSRTs. SSRTs were estimated using the quantile method (Band et al., 2003). The only effect to emerge in the analysis was a main effect of Block Type [$F(1, 77) = 12.55$, $p < .001$, $\eta_p^2 = .14$], such that inhibition was faster for stereotypic ($M = 482$ ms, $SD = 105$ ms) compared to counter-stereotypic ($M = 529$ ms, $SD = 106$ ms) responses.⁹

Table 2. Task performance as a function of Judgment and Block Type (Experiment 2).

Measure	Occupation		Trait	
	Stereotypic	Counter-stereotypic	Stereotypic	Counter-stereotypic
Go RT (ms)	566 (10)	570 (11)	606 (11)	615 (12)
Accuracy (%)	96 (1)	94 (1)	96 (1)	95 (1)
SSD (ms)	263 (9)	253 (7)	275 (8)	269 (8)
Failed inhibition (%)	33 (3)	35 (2)	28 (2)	32 (2)
SSRT (ms)	252 (16)	264 (15)	254 (7)	274 (9)

Note. RT = reaction time. Standard error of the mean (SEM) in parentheses.

⁹ No association between Go RTs and SSRTs was observed, $r(77) = -0.079$, $p = .486$.

Discussion

Using a different paradigm and measure of stereotyping, these results replicate those observed in Experiment 1. Following the presentation of male and female faces, response inhibition (i.e., SSRTs) was faster for stereotypic compared to counter-stereotypic judgments, regardless of the stereotype-related dimension under consideration (i.e., occupational or personality-related information).

General Discussion

A widely endorsed viewpoint is that construing others in an expectancy-confirming (i.e., stereotypic) manner is the mind's default outcome, taking less time and effort than the expectancy-disconfirming (i.e., counter-stereotypic) alternative (Allport, 1954; Bar, 2004, 2007; Bodenhausen & Macrae, 1998; Freeman & Ambady, 2011; Macrae & Bodenhausen, 2000). An interesting consequence of this propensity is that inhibitory demands may be greater when counter-stereotypic compared to stereotypic responses must be suppressed (Logan, Van Zandt, Verbruggen, & Wagenmakers, 2014; Verbruggen et al., 2004, 2006). The current research yielded evidence for just such an effect. Countering the assumption that stereotypic responses are difficult to inhibit (Bargh, 1999; Fiske & Neuberg, 1990; Macrae & Bodenhausen, 2000), here we demonstrated quite the opposite. When stopping stereotype-related responses, SSRTs were faster for stereotypic compared to counter-stereotypic judgments. This effect, moreover, emerged whether stereotyping was probed indirectly (Expt. 1 – sex categorization) or directly (Expt. 2 – stereotype-based judgments).

That stereotypic rather than counter-stereotypic responses are easier to inhibit is consistent with related work using the flanker task (Eriksen & Eriksen, 1974). In the flanker task, participants make speeded responses to stimuli (e.g., letters) which are flanked by distractors. Critically, these distractors are either congruent (i.e., require the same response) or incongruent (i.e., require a different response) with respect to the target stimulus. The widely replicated finding in this

paradigm is that responses are slower when targets are accompanied by incongruent compared to congruent flankers. Interestingly, SSRTs yield the same pattern, such that stopping is impaired when flankers are incongruent rather than congruent with the target. Ridderinkhoff et al. (1999) attribute this effect to the interaction between two inhibitory demands (Nigg, 2000), the requirement to inhibit an incorrect response in the flanker task and the suppression of a motor response when a stop-signal is presented (see also Kramer et al., 1994; Schachar et al., 2007; Verbruggen et al., 2004, 2006). A similar interpretation can be advanced for the current findings. On counter-stereotypic (vs. stereotypic) stop trials, participants must suppress both a motor response (i.e., behavioral inhibition) and the incorrect judgmental response (i.e., interference control) that has been triggered by the priming stimulus (Verbruggen et al., 2004, 2006). As a result, stopping is slowed on counter-stereotypic compared to stereotypic trials.

In interpreting the current effects, a caveat is in order. In tasks of the sort used here (e.g., Stroop, flanker), incongruent stimuli are known to hinder both the execution and inhibition of responses. That is, Go RTs and SSRTs are impaired by response incompatibility. As such, the precise origin of the inhibitory effects reported here remains open to question. To address this issue, one possibility would be to include a stereotype-neutral condition in which prior expectancies do not guide response selection. In Experiment 1, for example, participants could have performed a sex-categorization task on faces that were preceded by a gender-neutral occupational context. It is worth noting, however, that previous research has revealed that stopping is more difficult on incongruent than congruent and neutral trials (Kramer et al., 1994), thereby tracing impairments in response inhibition to the combination of behavioral suppression and interference control during periods of response conflict (i.e., stopping performance is equivalent on congruent and neutral trials). Nevertheless, to clarify and extend the current findings, future research should explore the ease with which both stereotype-related (i.e., stereotype consistent/stereotype-inconsistent) and non-stereotypic (i.e., stereotype-neutral) responses can be inhibited. In so doing, it will be possible

to compare stereotyping research with work exploring the efficiency of response inhibition in other cognitive tasks.

The demonstration that stereotypic responses can be stopped quickly offers reassurance to those concerned about the controllability of this process (Bargh, 1999; Blair, 2002; Moskowitz, 2010). However, before concluding that the deleterious consequences of stereotyping can readily be subdued, a cautionary caveat is in order. Here we demonstrated the intentional inhibition of stereotype-related responses. Although stereotyping outside the laboratory can, and indeed often does, entail the deliberate suppression of discriminatory actions, worries about this practice frequently dwell on the implicit channels through which stereotypes bias behavior. For example, unbeknownst to an individual, stereotypes can subtly influence how information is perceived, interpreted, represented in memory, and ultimately used to guide responses toward other people (Freeman & Ambady, 2011; Macrae & Bodenhausen, 2000). Whilst it has been argued that top-down control can be exerted over seemingly impenetrable processing operations (Lupyan, 2015), the current findings do not speak to this matter. Instead, emphasis falls on executive control — specifically, the ability to implement goal-directed behavior on request (Diamond, 2013; Logan & Cowan, 1984). A useful task for future research will therefore be to consider how different inhibitory functions contribute to stereotype control (Nigg, 2000). In particular, in tandem with the controlled inhibitory operations that enable people to cancel stereotyped outputs on demand, when do automatic inhibitory processes restrain the influence of pre-potent stereotype-based responses (Schachar et al., 2007; Verbruggen et al., 2006). Work of this kind will be important as harmonious living is underpinned by both these forms of inhibition (Payne, 2005).

To extend the scope of the current investigation, other issues also merit empirical scrutiny. These center upon *how* stereotyping is measured, *where* it takes place, and *who* is responding in this way. Using standard methodologies to explore person construal (Freeman & Ambady, 2011; Macrae & Bodenhausen, 2000), the current research indexed the effects of expectancy-based

processing via the activation of gender stereotypes (Quadflieg et al., 2011). Interestingly, comparable effects emerged whether stereotype inhibition was probed indirectly (i.e., sex categorization) or directly (i.e., overt stereotype-based judgments). Although this approach was productive, future work should consider the ease with which stereotype-related responses can be inhibited for a wider range of commonly stereotyped groups. What seems likely is that stereotype inhibition may be sensitive to the strength of people's prior beliefs and the specific stereotypes under consideration (Bar, 2004, 2007). For example, differences in the extent to which stereotypes are endorsed, in combination with people's executive abilities and desire to appear non-prejudiced, may influence the ease with which associated behavioral outputs can be stopped (Moskowitz, 2010; Moskowitz, Gollwitzer, Wasel, & Schaal, 1999). In particular, whereas people are probably highly practiced at inhibiting culturally sensitive stereotypes, for other social groups the tendency to avoid stereotyping is unlikely to have been developed (Devine, 1989; Payne, 2005). Exploring this issue will advance understanding of how perceiver-related factors in combination with prevailing societal norms influence the efficiency of stereotype control.

Conclusion

Acknowledging the perils of stereotyping (Fiske, 1998), the current research considered the ease with which responses to stereotype-related stimuli can be stopped (Bartholow et al., 2006). Opposing conventional wisdom (Bargh, 1999), the results revealed enhanced performance when stereotypic rather than counter-stereotypic responses were inhibited. Of course, whether these results generalize to other measures and manifestations of stereotyping remains an important question for future research. Nevertheless, here we provide first evidence that stopping stereotypic responses may be less challenging than has previously been assumed.

Open Practices Statement

The data and materials are available on request from the first author (JKF). None of the experiments were preregistered.

References

- Allport, G.W. (1954). *The nature of prejudice*. Reading, MA: Addison- Wesley.
- Band, G. P. H., van der Molen, M., & Logan, G. D. (2003). Horse-race model simulations of the stop-signal procedure. *Acta Psychologica*, 112, 105-142.
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617-629.
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11, 280-289.
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY: Guilford Press.
- Bartholow, B. D., Dickter, C. L., & Sestir, M. A. (2006). Stereotype activation and control of race bias: Cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, 90, 272-287.
- Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: Control adjustments in the stop-signal paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 392-404.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6, 242-261.
- Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. *Advances in Social Cognition*, 11, 1-52.
- Brewer, M. B. (1988). A dual process model of impression formation. *Advances in Social Cognition*, 1, 1-36.

- Cebr (2018). The value of diversity. *An Involve report prepared by Cebr*. London, UK.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143-149.
- Falbén, J. K., Tsamadi, D., Golubickis, M., Olivier, J. L., Persson, L. M., Cunningham, W. A., & Macrae, C. N. (in press). Predictably confirmatory: The influence of stereotypes during decisional processing. *Quarterly Journal of Experimental Psychology*.
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (pp. 357-411). New York, NY: McGraw-Hill.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1-74.
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*, 247-279.
- Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101-135.

- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47, 237-271.
- 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, 199-227.
- Kramer, A. F., Humphrey, D. G., Larish, J. F., Logan, G. D., & Strayer, D. L. (1994). Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. *Psychology and Aging*, 9, 491-477.
- Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295-327.
- Logan, G. D., Cowan, W. B., & Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: A model and method. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 276-291.
- Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, 121, 66-95.
- Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology*, 6, 547-569.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47, 1122-1135.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51, 93-120.

- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*, 808-817.
- Monteith, M. J., Sherman, J. W., & Devine, P. G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology Review*, *2*, 63-82.
- Moskowitz, G. B. (2010). On the control over stereotype activation and stereotype inhibition. *Social and Personality Psychology Compass*, *4*, 140-158.
- Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology*, *77*, 167-184.
- Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*, 220-246.
- OECD (2016). The economic cost of gender-based discrimination in social institutions. *OECD Development Centre*. OECD Publishing, Paris.
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition*, *25*, 273-285.
- Payne, B. K. (2005). Conceptualizing control in social cognition: How executive functioning modulates the expression of automatic stereotyping. *Journal of Personality and Social Psychology*, *89*, 488-503.
- Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., & Macrae, C. N. (2011). Stereotype-based modulation of person perception. *Neuroimage*, *57*, 549-557.

- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260-281.
- Ridderinkhof, K. R., Band, G. P. H., & Logan, G. D. (1999). A study of adaptive behavior: Effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychologica*, *101*, 315-337.
- Schachar, R., Logan, G. D., Robaey, P., Chen, S., Ickowicz, A., & Barr, C. (2007). Restraint and cancellation: Multiple inhibition deficits in attention deficit hyperactivity disorder. *Journal of Abnormal and Child Psychology*, *35*, 229-238.
- Verbruggen, F., Liefoghe, B., & Vandierendonck, A. (2004). The interaction between stop signal inhibition and distractor interference in the flanker and Stroop task. *Acta Psychologica*, *116*, 21-37.
- Verbruggen, F., Liefoghe, B., & Vandierendonck, A. (2006). The effect of interference in the early processing stages on response inhibition in the stop signal task. *Quarterly Journal of Experimental Psychology*, *59*, 190-203.
- Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*, 418-424.
- Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigm. *Neuroscience and Biobehavioral Reviews*, *33*, 647-661.
- Wang, P., Tan, C-H., Zhang, Y. L. Q., Wang, Y-B., & Luo, J-L. (2018). Event-related potential N270 as an index of social information conflict in explicit processing. *International Journal of Psychophysiology*, *123*, 199-206.
- Wentura, D., & Rothermund, K. (2014). Priming is not priming is not priming. *Social Cognition*, *32*, 47-67.

- White, C. N., Congdon, E., Mumford, J. A., Karlsgodt, K. H., Sabb, F. W., Freimer, N. B., London, E. D., Cannon, T. D., Bilder, R. M., & Poldrack, R.A. (2014). Decomposing decision components in the stop-signal task: A model-based approach to individual differences in inhibitory control. *Journal of Cognitive Neuroscience*, *26*, 1601-1614.
- White, K. R., Crites Jr, S. L., Taylor, J. H., & Corral, G. (2009). Wait, what? Assessing stereotype incongruities using the N400 ERP component. *Social Cognitive and Affective Neuroscience*, *4*, 191-198.
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: hierarchical Bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, *7*, 14.
<http://doi.org/10.3389/fninf.2013.00014>.