Cognitive, Affective, & Behavioral Neuroscience (2018) 18:1283–1297
https://doi.org/10.3758/s13415-018-0639-8

Neural and behavioral effects of regulating emotional responses to errors during an implicit racial bias task

Meredith P. Levsen[1] **· Bruce D. Bartholow**[1]

Abstract

Affect regulation plays a key role in several theories of racial bias reduction. Here, we tested whether engaging in emotion regulation strategies while performing an implicit racial bias task (Weapons Identification Task; WIT) would alter neural and behavioral manifestations of bias. Participants either suppressed or reappraised in a positive light the distress associated with making errors during the WIT, while an electroencephalogram (EEG) was recorded. We hypothesized that engaging in emotion regulation strategies would reduce the distress associated with making errors indicative of bias, resulting in smaller error-related negativity (ERN) amplitude during errors and increased expression of racial bias. Results of within-subjects comparisons (Experiment 1) generally supported these predictions. However, when emotion regulation strategies were manipulated between subjects (Experiment 2) there was no effect of suppression or reappraisal on bias expression. Across both experiments, engaging in emotion regulation led to larger ERNs for errors occurring on Black- relative to White-primed trials. In addition, a number of significant order effects were observed, indicating important differences in the effects of engaging in emotion regulation strategies when those strategies are attempted in participants' first versus second block of trials. No such order effects were evident when a second trial block was completed with no emotion regulation instructions. Findings are discussed in terms of the need for greater specificity in experimental tests of emotion regulation on error processing and cognitive performance.

Keywords ERN · Emotion · Regulation · Control · Racial bias

Introduction

Actions often fall short of intended goals, producing an acute state of conflict that, according to the conflict monitoring theory (see Botvinick et al., 2001; Carter & van Veen, 2007), manifests in the brain via heightened activation of the dorsal anterior cingulate cortex (dACC) and other structures in the so-called salience network (e.g., Ham, Leff, de Boissezon, Joffe, & Sharp, 2013). Depending on the personal significance of the behavior-intention discrepancy, such conflict may be experienced as aversive, as indicated by evidence from self-report (Spunt, Lieberman, Cohen, & Eisenberger, 2012), behavioral (Aarts, De Houwer, & Pourtois, 2013), and psychophysiological measures (Bartholow, Henry, Lust, Saults, & Wood, 2012; Braem et al., 2017; Hajcak & Foti, 2008; Lieberman & Eisenberger, 2015; Lindstrom, Mattsson-Marn, Golkar, & Olsson, 2013; Pourtois et al., 2010). Moreover, conflict-related aversive affect may be instrumental for engaging cognitive control (Cavanagh & Shackman, 2015; Dreisbach & Fischer, 2015; Inzlicht, Bartholow, & Hirsch, 2015). The purpose of the current research was to investigate the implications of attempts to regulate the affect associated with error commission for neural and behavioral manifestations of control, in the specific context of implicit racial bias.

Applied to understanding expression of racial bias, conflict monitoring can alert people to instances when their behavior falls short of their goal to be egalitarian. Amodio et al. (2004) first demonstrated that errors indicative of bias in a laboratory task are accompanied by a scalp-recorded reflection of dACC activation, the error-related negativity (ERN; see Debener et al., 2005; van Veen & Carter, 2002), and that variability in the magnitude of this "race-bias ERN" predicts control of bias in the task. Such findings (also see Amodio, Kubota, Harmon-

---

Electronic supplementary material The online version of this article (https://doi.org/10.3758/s13415-018-0639-8) contains supplementary material, which is available to authorized users.

✉ Bruce D. Bartholow
bartholowb@missouri.edu

[1] Department of Psychology, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, USA

Jones, & Devine, 2006; Amodio, Harmon-Jones, & Devine, 2008) suggest that the experience of compunction associated with a failure to control implicit bias (see Devine, Monteith, Zuwerink, & Elliot, 1991; Monteith, 1993) contributes to the control of bias expression overall (Monteith, Ashburn-Nardo, Voils, & Czopp, 2002; Monteith, Mark, & Ashburn-Nardo, 2010), and that the dACC plays a key role in this process (see Bartholow et al., 2012; Inzlicht et al., 2015).

To the extent that error-related distress is important for controlling bias expression, it could be that attempts to modify that affective experience will have implications for both error-related dACC activation and behavioral indices of control. Deliberate attempts to alter the experience or expression of affect have been termed *emotion regulation* (Gross, 1998). *Cognitive reappraisal*, a so-called antecedent-focused strategy because it can be implemented prior to the emergence of an emotion, involves reframing the affect-eliciting stimulus, or thinking about it from a different perspective, in order to alter its meaning (e.g., Ochsner, Silvers, & Buhle, 2012). *Expressive suppression*, in contrast, is a response-focused strategy involving inhibition of the external behavioral signs of an emotional response (Gross, 2001, 2002; Gross & John, 2003; Gross & Levenson, 1993; Butler, Lee, & Gross, 2007).

Recently, researchers have begun to investigate whether implementing emotion regulation can affect dACC activation associated with errors. In an fMRI study, Ichikawa et al. (2011) instructed participants to down-regulate their emotional responses to errors during a challenging cognitive task by using a mixture of regulation strategies. These authors reported no effects of emotion regulation strategy in ROI analyses focused on the ACC's response to errors, and observed predicted effects of emotion regulation on dACC activation (i.e., decrease < natural) only very late during error trials (~9 s post-response). In another study, Hobson, Saunders, Al-Khindi, and Inzlicht (2014) told participants to "adopt a detached attitude" and "view the task from a third-person perspective" (p. 1017; i.e., psychological distancing; see Shiota & Levenson, 2009, 2012), and found that this strategy was effective in reducing ERN amplitude during response errors relative to a no-instruction control condition. Furthermore, participants' ratings of emotional involvement (given following each condition) mediated the association between instructions and changes in the ERN.

These two studies converge to suggest that activation of the dACC during errors can be modified by applying versions of classic emotion regulation strategies. However, a number of questions remain concerning potential differences in the effects of different "down-regulation" strategies on the ERN, as well as whether any such effects have implications for the control and expression of race bias. For example, participants in both Ichikawa et al. (2011) and Hobson et al. (2014) completed the tasks under each set of emotion-regulation instructions. Given the potential for carryover or demand characteristics to affect responses under different instruction conditions, characterizing effects measured under within- and between-participants designs is important to establish the boundaries of these effects.

The purpose of the current studies was to investigate effects of implementing reappraisal and suppression strategies on the control and expression of racial bias during the Weapons Identification Task (WIT; Payne, 2001), and on the amplitude of the ERN elicited by failures of control. Based on previous research suggesting that emotion regulation strategies can reduce dACC reactivity to errors (Hobson et al., 2014; Ichikawa et al., 2011), and on the more general notion that reappraisal can decrease negative affect (e.g., Ochsner et al., 2012), we predicted that participants engaging in reappraisal would show reduced ERN amplitudes, especially for errors indicating racial bias.[1] Following the logic that error-induced negative affect is important for implementing control following errors (Bartholow et al., 2012; Inzlicht et al., 2015), and that dACC/ERN amplitude reflects the magnitude of such affective responses (e.g., Eisenberger, 2012), we further predicted that utilizing reappraisal would reduce the influence of control-related processes on performance, as determined by process dissociation procedure (PDP) analyses (see Jacoby, 1991; Payne, 2001). In contrast to reappraisal, expressive suppression generally is considered less effective in regulating negative affective experience (e.g., Ochsner et al., 2012) and its neural correlates (see Goldin et al., 2008). Thus, we predicted that ERN amplitude would be less affected by suppression compared with reappraisal, and subsequently that suppression would not reduce the influence of control-related processes on performance.

These hypotheses were tested in two studies during which participants completed multiple blocks of the WIT under differing emotion-regulation instructions. In the first experiment all participants completed a control "attend" block first, to establish a baseline against which to compare effects of reappraisal and suppression instructions. In the second experiment, to permit between-subjects comparisons of the effects of all three instruction sets, a different sample of participants completed the WIT under either suppression or reappraisal instructions before completing an "attend" block.

---

[1] The ERN unfolds rapidly during commission of an error, which might make it seem difficult for reappraisal to affect ERN amplitude. However, as mentioned previously, reappraisal is considered an antecedent-focused emotion regulation strategy, taking place *before* the emotional response has completely materialized (Gross, 1998). To the extent that the ERN reflects an affective response to errors, and if reappraisal alters the materialization of that affective response, then we would expect reappraisal to modulate the ERN, despite the short time-frame. Moreover, we assume that the reappraisal instructions participants received in the study encourage a "reappraisal mindset," which should reduce the impact of errors on the generation of the ERN in a proactive manner.

General method

Overview

Participants in both experiments completed the WIT (Payne, 2001) while an electroencephalogram (EEG) was recorded. Stimuli, experimental task, and instructions concerning how participants should attempt to regulate their affect were identical across the two experiments; they differed only in the order in which participants completed each task block (i.e., instruction set).

Stimuli and experimental task

The WIT assesses racial bias stemming from the stereotype associating young Black men with gun violence. Each trial begins with a visual "white noise" mask shown for 1,000 ms. Then, a gray-scale picture of either a White or a Black male's face (prime) is shown for 200 ms, followed immediately by a gray-scale picture of either a handgun or a tool (target) also displayed for 200 ms, after which a second masking pattern is shown for 300 ms. Participants' task is to categorize each target as a gun or a tool as quickly as possible by pressing one of two buttons. Responses made following a 500-ms response deadline elicit a "Too Slow!" message; no feedback was provided for responses given before the deadline, because previous research has shown that participants are highly aware of the accuracy of their responses during the WIT even when no feedback is provided (see Bartholow et al., 2012; Payne, Shimizu, & Jacoby, 2005). Trials are separated by a randomly varying inter-trial interval of 800, 1,000, or 1,200 ms, during which a fixation cross is displayed. Racial bias is revealed by more accurate categorization of guns than tools on Black-primed trials (see Amodio et al., 2004, 2008; Ito et al., 2015; Payne, 2001, 2005). Each block consisted of 192 trials (48 of each of four prime-target pairings).

Emotion regulation instructions

The primary experimental manipulation in the current studies consisted of three sets of instructions (i.e., blocks) pertaining to regulation of the affect accompanying errors. In the *Attend* (i.e., control) block, participants were given standard WIT instructions. They were told that on each trial a face would appear briefly on the computer monitor, followed by a picture of either a handgun or a hand tool, and that their task was to classify the second picture as a gun or a tool as quickly and accurately as possible. In addition, all participants were told that "the task is designed to measure the racial prejudice of individuals." Prior to receiving instructions for the first emotion regulation block, the experimenter acknowledged that errors in the task can have an affective component:

> "During this task, people tend to feel badly when they make errors. It usually doesn't feel good to make an error in general, but in this context the additional racial bias component may cause people to feel even worse. For example, if you accidentally classify a tool as a gun after seeing a Black face, this can be indicative of unconsciously endorsing Black stereotypes."

This was conveyed to participants in order to first acknowledge and identify the emotions they would be asked to regulate, before the experimenter could explain the strategy they would be using to do so.

For the *Suppression* block, participants were further instructed to not express any externally visible, behavioral signs of emotion whenever they made errors (see Bebko et al., 2011; Gross & Levenson, 1993):

> "During this block, whenever you make an error, you should to try to suppress your feelings about it. What this means is that you're going to try to not externally express any thoughts or feelings that come to mind about the error. You can think about it like trying to have a 'poker face' – if someone else was in the room with you, they wouldn't be able to tell what kinds of emotions you were experiencing."

For the *Reappraisal* block, participants were instructed to "regulate" their emotions by doing the following (see Ochsner et al., 2002):

> "During this block, whenever you make an error, try to think about it in such a way that you feel less negative about it. For example, you can tell yourself that everyone makes mistakes—it's really not a big deal. You can also view each error as an opportunity to learn and do better, knowing there are many, many trials and plenty of opportunities to keep trying."

Electrophysiological recording

The EEG was recorded using 24 Ag/AgCl electrodes fixed in a stretch-lycra cap (ElectroCap, Eaton, OH, USA) placed on the scalp in standard locations (American Encephalographic Society, 1994). Scalp electrodes were referenced online to the right mastoid and an average mastoid reference was calculated offline. Vertical and horizontal eye movements were recorded with additional electrodes placed 1 cm above and below the left eye and approximately 2 cm outside the outer canthus of each eye, respectively. Electrode impedance was kept below 8 kΩ at all locations. EEG signals were amplified with Synamps2 amplifiers (Compumedics-Neuroscan, Charlotte, NC, USA), sampled at 500 Hz and filtered online at 0.01–40 Hz. Ocular artifacts were removed offline using a regression-based procedure (Semlitsch, Anderer, Schuster, & Presslich, 1986). A band-pass filter of 1–15 Hz was applied prior to deriving response-locked epochs of 1,000 ms (400 ms pre-response baseline). Following rejection of trials containing movement and other artifacts, EEG data were averaged according to experimental conditions for each electrode and participant.

General procedure

After obtaining informed consent, an experimenter led participants to an acoustically shielded room for placement of EEG electrodes. After receiving instructions and completing eight practice trials, participants performed the WIT while EEG was recorded. Each WIT block took approximately 10 min to complete. After the final WIT block the experimenter removed the electrode cap and showed participants to a private restroom where they could clean the electrode gel from their face and hair. Finally, participants completed a packet of questionnaires (see Supplemental Material) prior to being debriefed and excused.

Analytic approach

Behavioral data Response accuracy was scored by calculating the proportion of correct responses (excluding missed trials) in each condition. Accuracy rates were submitted to a 3 (Block instructions: Attend, Suppression, Reappraisal) × 2 (Prime: Black, White) × 2 (Target: gun, tool) analysis of variance (ANOVA). In Experiment 1, block instructions varied within subjects; in Experiment 2, block instructions varied between subjects. Response accuracy data also were used to calculate estimates of the influence of automatic and controlled processes using PDP equations (see Jacoby, 1991). PDP estimates of controlled (PDP-C) and automatic (PDP-A) processing were submitted to separate 3 (Block Instructions) × 2 (Prime) ANOVAs.[2] When appropriate, 95% confidence intervals (CIs) for estimates are reported.

ERP data Visual inspection of the response-locked EEG waveforms showed a pronounced negativity emerging approximately 30–110 ms following response onset and most prominent over frontal and fronto-central scalp sites (see Fig. 1), which we identified as the ERN. ERN amplitude was measured from a 3 × 3 matrix of fronto-central electrodes centering on FCz, where the ERN was largest. For each participant, only conditions containing at least five artifact-free error trials were included in data analyses ($M = 15.31$ [$SD = 7.29$] and $M = 12.72$ [$SD = 6.21$] trials were included in each participants' averages for Experiments 1 and 2, respectively).[3] Internal consistency of the ERN was good (αs ranged from .85 to .87 across Experiments 1 and 2; see Thigpen, Kappenman, & Keil, 2017).

However, this exclusion criterion resulted in a missing data problem, whereby 57% of participants across Experiments 1 and 2 were missing ERN data in at least one cell of the design. Repeated-measures ANOVA uses listwise deletion of cases with missing values, which would have resulted in elimination of these participants. Fortunately, we were able to use all possible observations by analyzing ERN amplitudes using multilevel modeling (MLM) with restricted maximum likelihood estimation (SAS 9.3; SAS Institute, Cary, NC, USA). In addition to accommodating differing patterns of missing observations across individuals, MLM has several other advantages over repeated-measures ANOVA for analyzing psychophysiological data (for in-depth consideration of these issues, see Bagiella, Sloan, & Heitjan, 2000; Kristjansson, Kircher, & Webb, 2007; Page-Gould, 2017). A variance-components covariance structure was used to estimate random intercept coefficients for participants and electrodes nested within participants. Degrees of freedom were calculated using the containment method (see West, Welch, & Galecki, 2015). Planned comparisons were conducted by comparing the estimated marginal population means via paired *t*-tests (calculated with the LSMEANS /DIFF statement in SAS). All reported ERN means are the model-specific estimated marginal means (in μV). Average ERN amplitudes across conditions were analyzed using a multilevel modeling approach with random intercepts for subject and electrode nested within subject.[4] Effect sizes were calculated via the partial R^2 method (Edwards, Muller, Wolfinger, Qadish, & Schabenberger, 2008), where $R^2 < .1304$ is considered a small effect, $.1304 \leq R^2 < .2592$ is a medium effect, and $R^2 \geq .2592$ is a large effect (Cohen, 1992).

---

[2] As explained by Payne (2005), the PDP approach assumes that all behavior is determined by a combination of automatic and controlled processes. The relative influence of these processes can be estimated in tasks in which some trials allow these processes to act in concert (i.e., *congruent trials*, such as Black-gun) while others place these processes in opposition (i.e., *incongruent trials*, such as Black-tool). For each participant, the control estimate *C* is computed as the proportion of congruent trials on which they responded correctly minus the proportion of incongruent trials on which they committed a stereotype-related error (e.g., responding with the "gun" key on black-tool trials). The automatic estimate *A* is that same proportion of incongruent error trials divided by the inverse of *C*. (The full set of PDP equations can be found in Payne, 2005.)

[3] The number of trials contributing to each ERN average (per each subject × Block Instructions cell), broken down by WIT conditions across both experiments, was as follows: for Black-gun, $M = 11.38$, $SD = 5.48$; for Black-tool, $M = 16.60$, $SD = 7.64$; for White-gun, $M = 15.42$, $SD = 6.77$; for White-tool, $M = 13.01$, $SD = 6.52$. For an additional breakdown of Block Instructions by Experiment, see Table S2 (Supplemental Material).

[4] Including multiple channels and specifying electrode as a random factor extracts variance in common across the channels that is related to the model's predictors. Residual variance related to electrodes is partitioned separately from other sources of variance, and therefore it does not unduly increase the standard error of the estimates for the fixed predictors. This, in turn, increases statistical power to detect effects associated with the predictors (Tibon & Levy, 2015; Vossen et al., 2011).

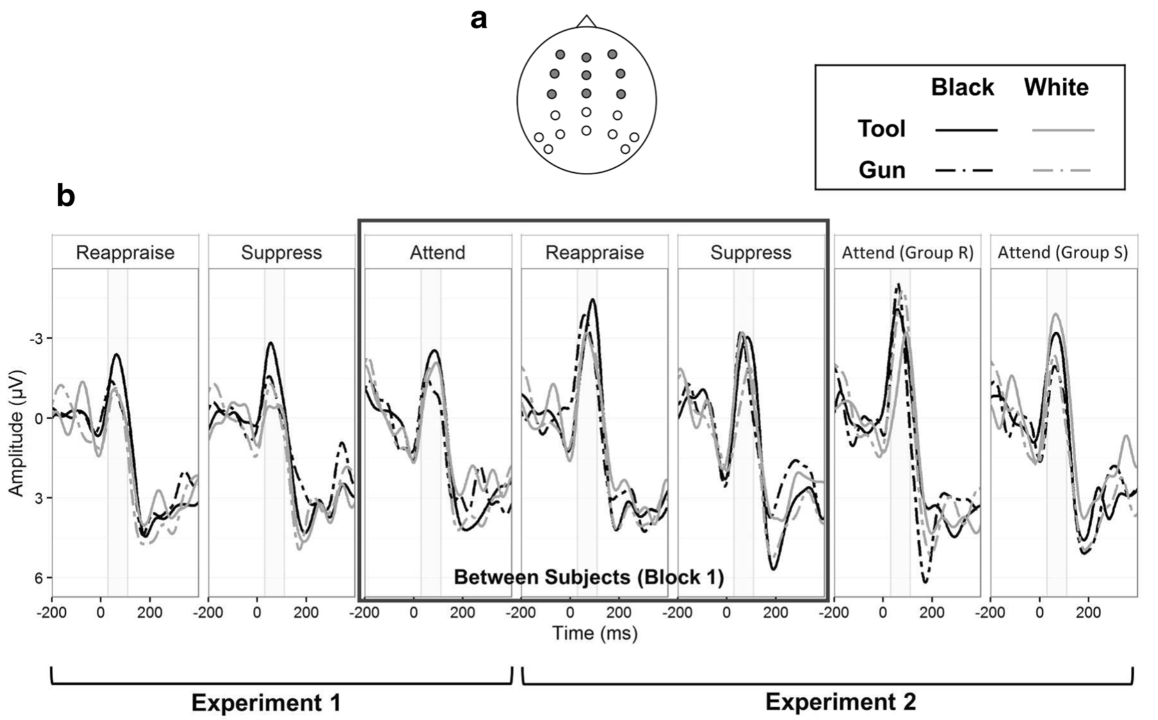

Fig. 1 **Panel A**: Electrode montage used for EEG recording, highlighting electrodes (gray) included in the ERN analyses (F3, Fz, F4, FC3, FCz, FC4, C3, Cz, and C4). **Panel B**: Grand average, response-locked ERP waveforms elicited during errors as a function of prime race, target type, and emotion-regulation instructions. Groups of waveforms enclosed in the box are those used in between-subjects comparisons of first-block performance. "Attend (Group R)" and "Attend (Group S)" indicate the Attend blocks for participants who first completed a Reappraisal block and a Suppression block, respectively, in Experiment 2. Grand averages were weighted by the number of trials and subjects per condition. Shading indicates the time interval during which ERN amplitudes were quantified (30–110 ms post-response)

Experiment 1

Method

Participants

We aimed to have at least 40 participants with valid data – this sample size is equal to or greater than sample sizes used in previous experiments examining the effects of emotion regulation on error processing (*N*s = 41 and 17 in Hobson et al., 2014 and Ichikawa et al., 2011, respectively). Forty-seven undergraduates (M_{age} = 19 years; *n* = 20 women) at the University of Missouri volunteered in exchange for partial course credit. The sample was 6% Hispanic, 9% Black, 9% Asian, 4% multiracial, and 72% White/Non-Hispanic. All participants met basic eligibility criteria for participation in an EEG study (no history of head trauma or neurologic disease; no dreadlocks, cornrows or braided hair).

Design

All participants completed the WIT under each instruction condition, resulting in a 3 (Block Instructions; Attend, Suppression, Reappraisal) × 2 (Prime; Black, White) × 2 (Target; tool, gun) within-subjects design. To establish baseline measures of ERN amplitude and WIT performance and prevent carry-over effects from the emotion regulation instructions, all participants completed the WIT under Attend instructions first. The emotion regulation instructions were given prior to the second and third blocks; the order of these blocks was counter-balanced across participants, such that 24 completed the Suppression block before the Reappraisal block, and 23 completed the Reappraisal block before the Suppression block.

Results and discussion

Data from one of the three blocks were excluded for two participants (one used only one button to respond on all trials in the last block; one mistook the power drill tool for a gun during the first block). Thus, degrees of freedom vary across comparisons.

Response accuracy

The ANOVA on accuracy rates showed the typical Prime × Target interaction, $F(1, 44) = 80.51$, $p < .001$, $\eta_p^2 = 0.65$ (see Fig. 2, Table 1). Compared to White-prime trials, on Black-prime trials participants were less accurate at identifying tools ($t[44] = -7.39$, $p < .001$, CI [-.16, -.09]) and more accurate at

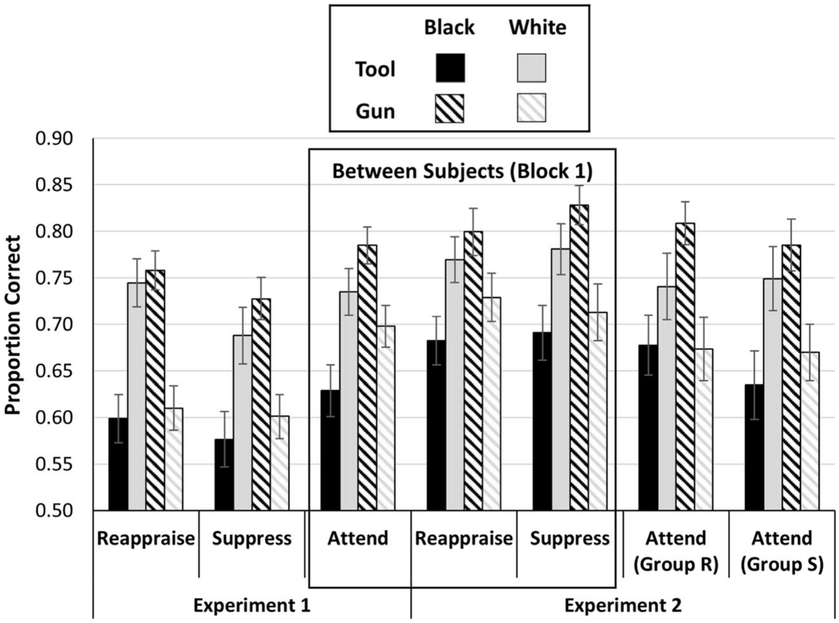

Fig. 2 Response accuracy as a function of emotion-regulation instructions. Groups of means enclosed in the box are those used in between-subjects comparisons of first-block performance. Reappraise and Suppress = Reappraisal and Suppression instructions, respectively. "Attend (Group R)" and "Attend (Group S)" indicate the Attend blocks for participants who first completed a Reappraisal block and a Suppression block, respectively, in Experiment 2. Capped bars indicate standard error of the mean

identifying guns ($t[44] = 9.10, p < .001$, CI [.09, .15]), consistent with previous research (Payne, 2001, 2005; Amodio et al., 2004, 2008). The ANOVA also yielded a main effect for block instructions, $F(2, 88) = 16.64, p < .001, \eta_p^2 = 0.27$. Planned pairwise comparisons revealed that accuracy during Attend was greater than during Reappraisal, $t(44) = 3.23, p = .002$, CI [.01, .06], both of which were greater than during Suppression, $ts(45) = -5.04$ and -2.99, $ps \leq .004$, CIs [-.09, -.04], [-.05, -.01], respectively.

The prediction that reappraisal (relative to Attend) would be associated with a more biased pattern of responding was tested with the Block instructions × Prime × Target

Table 1 Mean accuracy and process dissociation estimates from both experiments

	Block 1				Block 2/3			
	Black		White		Black		White	
	Guns	Tools	Guns	Tools	Guns	Tools	Guns	Tools
Attend								
PDP-C	0.41 (0.27)		0.44 (0.25)		0.46 (0.28)		0.42 (0.30)	
PDP-A	0.63 (0.14)		0.55 (0.15)		0.63 (0.17)		0.59 (0.19)	
Accuracy	.78 (.14)	.63 (.19)	.70 (.15)	.74 (.17)	.80 (.14)	.66 (.19)	.67 (.18)	.75 (.20)
Suppression								
PDP-C	0.52 (0.25)		0.49 (0.25)		0.30 (0.28)		0.28 (0.27)	
PDP-A	0.65 (0.15)		0.56 (0.16)		0.62 (0.14)		0.59 (0.16)	
Accuracy	.83 (.12)	.69 (.16)	.71 (.17)	.78 (.15)	.73 (.16)	.57 (.18)	.60 (.15)	.69 (.19)
Reappraisal								
PDP-C	0.49 (0.26)		0.50 (0.25)		0.35 (0.26)		0.35 (0.26)	
PDP-A	0.63 (0.19)		0.54 (0.20)		0.62 (0.14)		0.63 (0.16)	
Accuracy	.80 (.15)	.68 (.15)	.73 (.16)	.77 (.14)	.76 (.14)	.59 (.18)	.61 (.16)	.74 (.17)

Values in parentheses are standard deviations. Block 2/3 refers to the fact that means for reappraisal and suppression from the first experiment reflect measures derived from both the second and third blocks. Block 1 Attend values and Block 2/3 Suppression and Reappraisal values are from Experiment 1; the remaining values are from Experiment 2. Note that analyses of between-subjects effects described under Experiment 2 in the text include Attend values measured during Experiment 1 interaction. This interaction was significant, $F(2, 88) = 4.21$, $p = .018$, $\eta_p^2 = 0.09$. Inspection of the means in Table 1 suggests that this interaction was driven by a more biased pattern of responding in both emotion-regulation blocks compared to the Attend block. To directly test this possibility, we calculated an overall bias score for each block, defined as the additive effect of prime race on accuracy for both types of targets:

WIT bias = (White-tool – Black-tool) + (Black-gun – White gun)

Mean bias scores were greater in the Reappraisal ($M = .30$) than in the Attend block ($M = 0.19$), $t(44) = 2.95$, $p = .005$, CI [.03, .18], and were directionally (but not significantly) greater in Reappraisal compared to Suppression ($M = .24$), $t(44) = 1.77$, $p = .083$, CI [-.01, .12]. WIT bias in the Suppression and Attend blocks did not differ ($p = .234$, CI [-.13, .03]).

We also examined bias scores separately for tool targets and gun targets. For tools, the Reappraisal block ($M = .15$) was associated with marginally more bias than both the Attend ($M = .11$), $t(44) = 1.76$, $p = .085$, CI [-.01, .08] and Suppression blocks ($M = .12$), $t(44) = 2.01$, $p = .050$, CI [0, .07], which did not differ ($p = .95$, CI [-.05, .05]). For gun targets the differences were more pronounced, with Attend being associated with less bias ($M = .08$) than both Reappraisal ($M = .15$), $t(44) = -3.20$, $p = .003$, CI [-.11, -.03], and Suppression ($M = .13$), $t(44) = -2.09$, $p = .043$, CI [-.10, -.002], which did not differ ($p = .342$, CI [-.02, .07]). These patterns largely support the prediction that engaging in reappraisal would increase expression of bias.

Automatic and controlled processes

We predicted that whereas utilizing reappraisal (relative to Attend) would reduce the influence of control-related processes on performance, utilizing suppression would not. To test this prediction, PDP-C and PDP-A scores were analyzed in separate 3 (Block instructions) × 2 (Prime) repeated measures ANOVAs (see Table 1 and Fig. 3). The ANOVA on PDP-C scores showed only a main effect of block instructions, $F(2, 90) = 16.64$, $p < .001$, $\eta_p^2 = 0.27$. Consistent with predictions, PDP-C was greater during Attend ($M = 42.5$) than during Reappraisal ($M = .35$). However, contrary to predictions, PDP-C was also greater during Attend than during Suppression ($M = .29$). Given that PDP-C is derived from accuracy, when collapsed across the Prime race variable this main effect merely reflects the fact that accuracy was greater during Attend than during Reappraisal, both of which were greater than during Suppression.

Analysis of PDP-A revealed a marginal main effect of block instructions, $F(2, 90) = 2.79$, $p = .067$, $\eta_p^2 = 0.06$, qualified by a significant Block instructions × Prime interaction, $F(2, 90) = 5.65$, $p = .005$, $\eta_p^2 = 0.11$. Pairwise comparisons indicated that PDP-A scores on Black-prime trials did not differ across blocks ($ts < 1$, $ps > .30$). However, for White-prime trials PDP-A scores were increased during Reappraisal compared to both Attend, $t(45) = 3.98$, $p < .001$, CI [.04, .13], and Suppression, $t(45) = 2.03$, $p = .049$, CI [.0002, .08], which did not differ ($t = -1.51$, $p = .14$, CI [-.04, -.09]). Moreover, whereas PDP-A was greater for Black-relative to White-prime trials under Attend, $t(46) = 2.77$, $p < .008$, CI [.02, .15], PDP-A scores did not differ as a function of prime race under Suppression or Reappraisal ($ts < 1$).

ERN amplitude

To test the hypothesis that engaging in reappraisal would reduce the ERN (relative to Attend), mean ERN amplitudes were submitted to a 3 (Block instructions) × 2 (Prime) × 2 (Target) MLM with random intercepts for subjects and electrode nested within subjects.[5] This analysis showed a main effect for block instructions, $F(2, 4165) = 28.85$, $p < .001$, $R^2 = .014$. As predicted, the ERN was larger overall during Attend ($M = -1.60\ \mu V$) than during Reappraisal ($M = -0.73\ \mu V$), $t(4165) = -4.27$, $p < .001$. Somewhat unexpectedly, the ERN also was reduced during Suppression ($M = -1.09\ \mu V$), $t(4165) = -7.58$, $p < .001$, and ERNs during Reappraisal and Suppression also differed from each other, $t(4165) = 3.13$, $p = .002$.

The main effect of block instructions was qualified by a Block Instructions × Prime interaction, $F(2, 4165) = 9.43$, $p < .001$, $R^2 = .005$. To simplify depiction of this interaction, Black-White prime difference scores were calculated in each condition (see Fig. 4). A series of simple effect tests (collapsing across target type) showed no effect of prime race on the ERN during the Attend block ($t < 1$, $p > .80$, CI [-0.29, 0.35]). However, during both emotion-regulation blocks, ERNs elicited on Black-prime trials were larger than those elicited on White-prime trials, $ts(4165) = -2.83$ and -6.06, $ps \leq .005$, CIs [-0.15, -0.80], [-0.65, -1.27], for Suppression and Reappraisal, respectively. Critically, within Black-prime trials, ERNs were larger during Attend compared to Reappraisal, $t(4165) = -2.34$, $p = .019$, CI [-0.68, -0.06], but not Suppression ($t = -1.57$, $p = .12$, CI [-0.57, 0.06]). Within White-prime trials, ERNs were larger during Attend compared to both Reappraisal, $t(4165) = -8.31$, $p < .001$, CI [-1.68, -1.04], and Suppression, $t(4165) = -4.43$, $p < .001$, CI [-1.09, -0.42]. These data suggest that, despite reducing the ERN overall, engaging in either emotion-regulation strategy is associated with larger ERNs on Black- relative to White-prime trials.

The analysis also showed main effects for Prime, $F(1, 4165) = 24.77$, $p < .001$, $R^2 = .006$, and Target, $F(1, 4165) = 16.52$, $p < .001$, $R^2 = .004$, qualified by a Prime × Target

---

[5] Analyses accounting for ERP responses elicited during correct trials (i.e., the correct response negativity, or CRN) on the amplitude of the ERN are given in the Supplemental Materials. Those analyses produced findings that, in large part, are consistent with the ERN analyses presented in the main text.

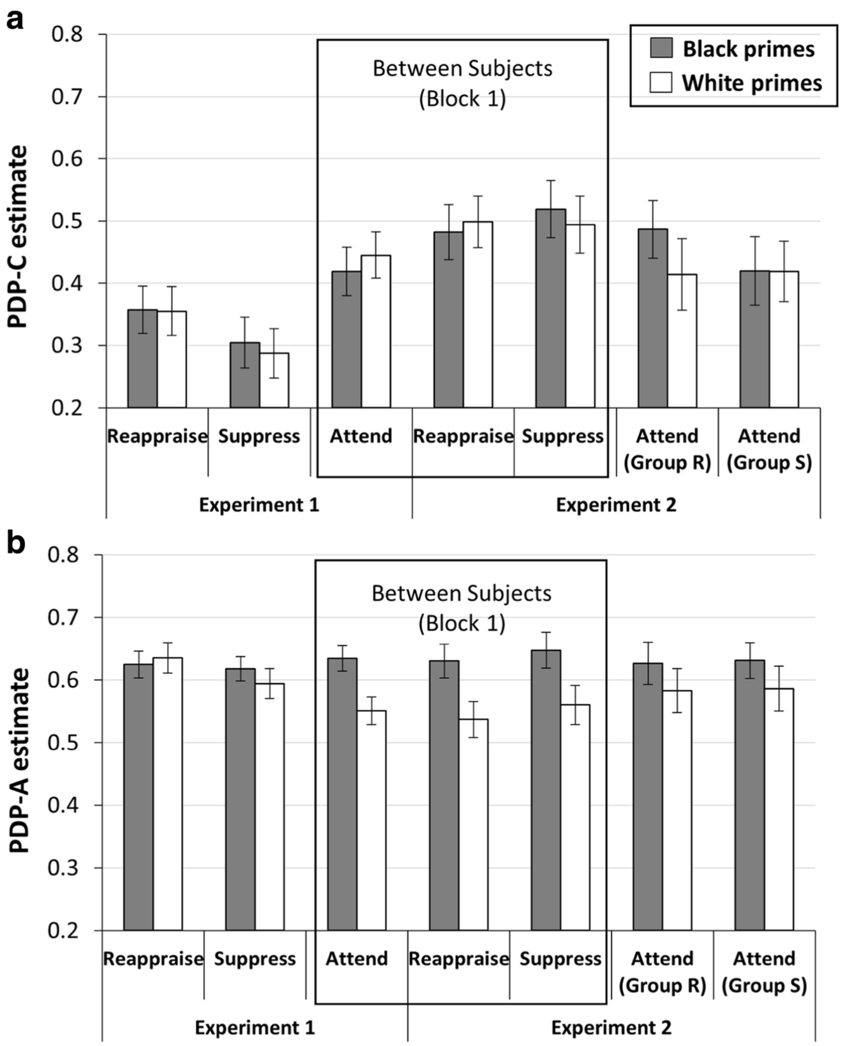

Fig. 3 Process-dissociation procedure estimates of controlled (**panel A**) and automatic processing (**panel B**) as a function of prime race and emotion-regulation instructions. Groups of means enclosed in the box are those used in between-subjects comparisons of first-block performance. Reappraise and Suppress = Reappraisal and Suppression instructions, respectively. "Attend (Group R)" and "Attend (Group S)" indicate the Attend blocks for participants who first completed a Reappraisal block and a Suppression block, respectively, in Experiment 2. Capped bars indicate standard error of the mean

interaction, $F(1, 4165) = 25.75$, $p < .001$, $R^2 = .006$. Collapsing across blocks, ERNs on Black-tool errors (M = -1.81, SE = 0.37, CI [-2.53, -1.08]) were larger than errors during all other trial types (Ms range from -0.86 to -0.95, SEs = 0.37, CIs range from -1.67 to -0.13), consistent with previous research (Amodio et al., 2004, 2008; Volpert-Esmond et al., 2018).

Overall, the pattern of results from Experiment 1 suggests support for the hypothesis that attempting to regulate the affect associated with errors leads to reduced error processing, poorer control, and increased response bias. Compared to the Attend block, bias was more pronounced (particularly for gun targets), PDP-C estimates were lower, and ERNs were smaller overall during the emotion-regulation blocks. Interestingly, ERNs elicited on Black-primed trials were larger than those elicited on White-primed trials during both emotion-regulation blocks, a pattern not seen during the Attend block.

The within-subjects design and block order used in Experiment 1 limits the inferences that can be drawn from the data. Given that completion of the WIT is thought to require cognitive control (see Ito et al., 2015; Payne, 2005), completing the Attend block may have strained participants' control-related resources. Thus, the observed reductions in accuracy, control, and ERN amplitude during the emotion-regulation blocks could be attributable to fatigue rather than (or in addition to) effects of emotion regulation per se. Similarly, attempting to implement an emotion-regulation strategy while performing a challenging cognitive task like the WIT likely also strains cognitive resources, and therefore increases in bias could be attributable to increased cognitive demand rather than reduction of error-related distress.

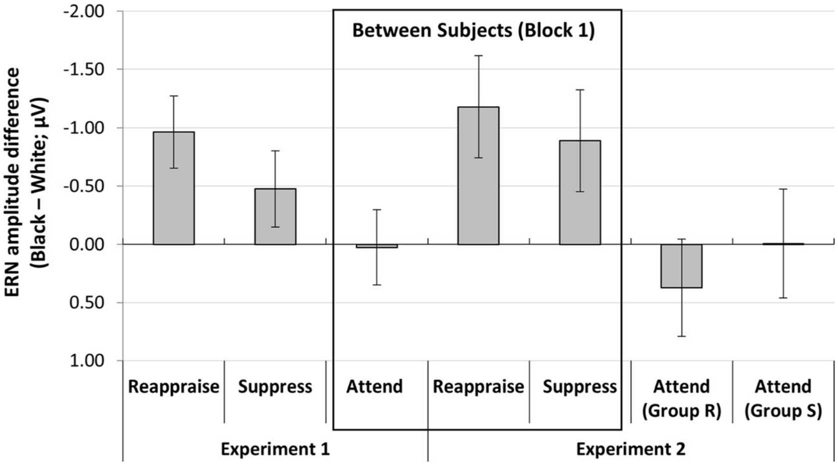

Fig. 4 Mean differences in ERN amplitude (estimated marginal means) elicited by Black-primed trials and White-primed trials as a function of emotion-regulation instructions. Means enclosed in the box are from the between-subjects multilevel model (MLM) comparisons of first-block performance. Means outside the box are derived from the within-subjects MLM from Experiment 1 (i.e., the two furthest left bars; Reappraise and Suppress) and Experiment 2 (i.e., the two furthest right bars; Attend [Group R] and [Group S]). Reappraise and Suppress = Reappraisal and Suppression instructions, respectively. Capped bars indicate 95% confidence intervals

Experiment 2 was designed to address these concerns. Participants completed the WIT under either Suppression or Reappraisal instructions first, followed by an Attend block. This design permits examination of the effects of emotion regulation uncontaminated by potential cognitive fatigue that might result from having previously completed the WIT. By combining data from the first (Attend) block of Experiment 1 with the first (Suppression or Reappraisal) block of Experiment 2, this design permits a between-subjects comparison of the effects of all three instruction sets when each is given in the first block. Concerns over fatigue were further reduced by having participants complete only two blocks (Suppression *or* Reappraisal, then Attend), which also eliminated the possibility of carry-over from one emotion-regulation block to another (cf. Hobson et al., 2014; Ichikawa et al., 2011).

Experiment 2

Method

Participants

Sixty-seven undergraduates participated in exchange for course credit in Introductory Psychology. Full data from two participants (and first-block data from another) were excluded due to a failure to understand and follow instructions. Data from the second (Attend) block were excluded for six participants due to difficulty discontinuing the regulation strategy, as indicated by their responses on a questionnaire administered following the experiment. These participants' data were included in the between-subjects analyses involving only the first block. EEG data from one participant were lost due to an equipment malfunction. The final sample used for first-block analyses included 64 participants (29 women; M_{age} = 18.6 years.), which achieved our goal of at least 30 participants per group for between-subjects analyses. This goal was based on previous neurophysiological studies of emotion regulation using between-subjects designs (max n = 26 per group; McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008; Opitz, Rauch, Terry, & Urry, 2012; van't Wout, Chang, & Sanfey, 2010). Roughly 11% of participants were Black, 8% were Asian, 2% were Hispanic, 5% were multiracial, and 75% were White.

Design and procedure

Participants were randomly assigned to one of two groups, which determined whether they engaged in Suppression or Reappraisal during their first trial block. Participants in both groups completed a second block under Attend instructions. Block instructions and task structure were identical to Experiment 1, but some minor procedural details were changed. Participants completed 16 rather than eight practice trials before hearing the emotion regulation instructions, and to reduce the potential for carryover effects participants were given a filler task after the first block (coloring an abstract, geometric design with crayons for 5 min). They were told that the purpose of this activity was to "to clear their mind" before being given different instructions for the second block.

Results and discussion

Analytic approach

Primary analyses focused on (between-subjects) comparison of behavioral and ERN responses measured during the first blocks of both experiments (Attend from Experiment 1 [n = 46]; Suppression [n = 30] and Reappraisal [n = 34] from Experiment 2). Data were analyzed using the same approaches as in Experiment 1, except that instruction condition was a between-subjects factor.

Although the main purpose of Experiment 2 was to compare effects of emotion-regulation instructions between subjects (to eliminate potential order effects in the data), the study design also permits examination of within-subjects effects for the two subject groups separately. To accomplish this, the accuracy data were subjected to separate 2 (Block; first block [Reappraise or Suppress], second block [Attend]) × 2 (Prime) × 2 (Target) repeated-measures ANOVAs, and the PDP-C and PDP-A data were subjected to separate 2 (Block) × 2 (Prime) repeated-measures ANOVAs.

Finally, it is also possible that block order effects differ according to instruction conditions. To examine this possibility, we submitted the data from the first two blocks of both experiments to a 2 (Block; block 1, block 2) × 3 (Instructions; Attend, Suppression, Reappraisal) mixed model, in which Block is a within-subjects variable (i.e., all participants completed a first and a second block), and Instructions varied both between subjects (i.e., not all participants completed Suppression or Reappraisal blocks) and within subjects (i.e., all subjects completed one Attend block and one emotion-regulation block). Therefore, this factor reflects a combination of both within- and between-subjects effects. Although repeated-measures ANOVA cannot accommodate this type of design, fortunately mixed modeling allows factors to represent a mixture of both between- and within-subjects effects.

Between-subjects comparisons of instruction conditions

Accuracy The ANOVA yielded a main effect for target, $F(1, 107) = 10.84$, $p = .001$, $\eta_p^2 = 0.09$, as well as the typical Prime × Target interaction, $F(1, 107) = 76.45$, $p < .001$, $\eta_p^2 = 0.42$ (see Table 2). In contrast to Experiment 1, this interaction was not further qualified by instruction condition, $F(2, 107) = 0.43$, $p = .651$, $\eta_p^2 = 0.01$.

Automatic and controlled processes The ANOVA on PDP-A scores showed a main effect of prime race, $F(1, 107) = 77.69$, $p < .001$, $\eta_p^2 = 0.42$; PDP-A was larger for Black- ($M = 0.64$, $SE = 0.02$, CI [0.61, 0.67]) than for White-primed trials ($M = 0.45$, $SE = 0.02$, CI [0.43, 0.48]). Contrary to Experiment 1, the Instructions × Prime interaction was not significant ($F < 1$), and the ANOVA on PDP-C scores did not show a main effect of instruction condition ($F < 1$).

ERN Our primary interest with the ERN analysis was in determining whether the Instructions × Prime interaction observed in Experiment 1 would replicate in the between-subjects design of Experiment 2. Indeed, the 3 (Instructions) × 2 (Prime) × 2 (Target) MLM revealed a significant Instructions × Prime interaction, $F(1, 2406) = 10.91$, $p < .003$, $R^2 = .005$. As in Experiment 1, errors on Black-primed trials elicited a larger ERN than errors on White-primed trials during both Reappraisal (Ms = -3.57 and -2.39, SEs = 0.64, CIs [-4.83, -2.30], [-3.65, -1.13], for Black and White, respectively, $t[2406] = -5.30$, $p < .001$, CI [-1.62, -0.74]), and Suppression (Ms = -2.15 and -1.26, SEs = 0.66, CIs [-3.45, -0.85], [-2.56, -.04] for Black and White, respectively, $t[2406] = -3.97$, $p < .001$, CI [-1.33, -0.45]).

The model also produced significant effects for both prime, $F(1, 2406) = 32.32$, $p < .001$, $R^2 = .013$, and target, $F(1, 2406) = 13.29$, $p < .001$, $R^2 = .005$. ERN amplitudes were greater for targets following Black primes ($M = -2.43$, $SE = 0.36$, CI [-3.13, -1.73]) than White primes ($M = -1.75$, $SE = 0.36$, CI [-2.45, -1.06]), $t(2406) = -5.69$, $p < .001$, CI [-0.91, -0.45], and were larger on tool errors ($M = -2.31$, $SE = 0.36$, CI [-3.01, -1.62]) than gun errors ($M = -1.87$, $SE = 0.36$, CI [-2.57, -1.17]), $t(2406) = -3.65$, $p < .001$, CI [-0.68, -0.20]. The Prime × Target interaction was not significant ($F < 1$).

These lower-order effects were qualified by a significant Instructions × Prime × Target interaction, $F(2, 2406) = 14.68$, $p < .001$, $R^2 = .012$. To unpack this complicated interaction, we examined the Prime × Target interaction separately within each level of instructions. For the Attend block, Prime × Target interaction was significant, $F(1, 1113) = 9.26$, $p = .002$, $R^2 = .008$. The ERN on Black-tool errors ($M = -2.03$, $SE = 0.49$, CI [-2.98, -1.07]) was larger than the ERNs elicited on all other trial types (Ms range from -1.13 to -1.70, SEs = 0.49, CIs max range [-2.65, -0.17]), $ts(1113) > -1.95$, $ps < .05$, CIs [-1.39, -0.41], [-1.03, -0.001]. The Prime × Target interaction was also significant for the Suppression block, $F(1, 639) = 21.53$, $p < .001$, $R^2 = .03$. Here, however, the pattern was different: ERNs during Black-tool, Black-gun and White-tool errors were equivalent (Ms range from -2.13 to -2.27, SEs = 0.76; $ts < 1$, $ps > .70$, Δs range from 0.03 to 0.14, CIs range [-0.55, 0.76]), and the ERN during White-gun errors was smaller than in those other conditions ($M = -0.25$, $SE = 0.76$). The ERN during White-gun errors differed from the ERN during White-tool, $t(654) = 6.66$, $p < .001$, CI [1.43, 2.62], and Black-gun trials, $t(654) = -6.23$, $p < .001$, CI [1.29, 2.47]. Finally, for the Reappraisal block, the Prime × Target interaction was not significant ($F < 1$, $p > .6$). These analyses strengthen the conclusion from Experiment 1 that utilizing Suppression or Reappraisal increases the ERN following Black face primes relative to White face primes.

Within-subjects comparison of instruction conditions

Accuracy The separate models of data from participants in the Reappraisal and Suppression groups both showed Prime × Target interactions, $F(1, 32) = 26.54$, $p < .001$, $\eta_p^2 = 0.45$; and $F(1, 29) = 38.99$, $p < .001$, $\eta_p^2 = 0.57$, respectively (see Table 1). These interactions reflect the typical pattern of WIT bias, where for Black- compared to White-prime trials participants made more errors when identifying tools ($t[32] = -4.00$, $p < .001$, CI [-0.11, -0.04]; $t[29] = -5.35$, $p < .001$, CI [-0.14, -0.06], for the reappraisal and suppression groups, respectively) and fewer errors when identifying guns, ($t[32] = 5.45$, $p < .001$, CI [0.06, 0.14]; $t[29] = 5.35$, $p < .001$, CI [0.07, 0.16], respectively). Additionally, participants were more accurate across groups when identifying guns compared to tools after seeing a Black face ($t[32] = -4.66$, $p < .001$, CI [-0.18, -0.07]; $t[29] = -5.49$, $p < .001$, CI [-0.20, -0.09], respectively).

The model examining data from the Reappraisal group also showed a significant Block × Prime interaction, $F(1, 32) = 8.44$, $p = .007$, $\eta_p^2 = 0.21$. Follow-up comparisons showed that, within the Attend block, accuracy was lower for White- ($M = .71$, $SE = .03$, CI [.65, .77]) compared to Black-prime trials ($M = .74$, $SE = .02$, CI [.70, .79]), $t(32) = 2.84$, $p = .008$, CI [0.01, 0.08]. However, within the Reappraisal block, accuracy on White-prime and Black-prime trials did not differ (*M*s = .75 and .74, *SE*s = .02, respectively), $t < 1$. Accuracy on White-primed trials also was lower during Attend than Reappraisal, $t(32) = 2.84$, $p = .008$, CI [0.01, 0.08].

The model examining data from the Suppression group showed a main effect of block, $F(1, 29) = 12.62$, $p = .001$, $\eta_p^2 = 0.30$, where participants were more accurate during the Suppression block ($M = .75$, $SE = 0.02$, CI [.71, .80] compared to the Attend block ($M = .71$, $SE = 0.03$, CI [.66, .76].

Automatic and controlled processes The ANOVAs examining PDP-A estimates showed no significant effects within either of the two participant groups, *F*s < 2.7, *p*s > .10.

The ANOVA examining PDP-C estimates for the Reappraisal group showed a significant Block × Prime interaction, $F(1, 32) = 8.44$, $p = .007$, $\eta_p^2 = 0.21$. Consistent with the accuracy data, this interaction was driven by lower control exhibited on White-prime trials ($M = .41$, CI [.30, .53]) relative to Black-prime trials ($M = .49$, CI [.39, .58]) during the Attend block, $t(32) = 3.42$, $p = .002$, CI [0.03, 0.11]. During the Reappraisal block, PDP-C did not differ for White-prime ($M = .50$, CI [.42, .59]) and Black-prime trials ($M = .49$, CI [.41, .58]), $t < 1$. PDP-C during White-prime trials also was lower during Attend than Reappraisal, $t(32) = 2.84$, $p = .008$, CI [0.02, 0.15].

For the Suppression group, the ANOVA on PDP-C estimates revealed only a main effect for block, $F(1, 29) = 12.62$, $p = .001$, $\eta_p^2 = 0.30$. PDP-C was higher during Suppression ($M = .51$, CI [.42, .60]) compared to Attend ($M = .42$, CI [.32, .52]). (*F*s < 1.1, *p*s > .3 for all other effects.)

ERN In some cases, the separate MLMs used to analyze the ERN data for the two groups garnered similar patterns and effects. For both models, there were main effects of prime, ($F[1, 1496] = 6.43$, $p = .011$, $R^2 = 0.004$; $F(1, 1347) = 6.36$, $p = .012$, $R^2 = 0.005$; within Reappraisal and Suppression groups, respectively). This effect was qualified in both models by a Block × Prime interaction ($F[1, 1496] = 25.51$, $p < .001$, $R^2 = 0.017$; $F[1, 1347] = 6.09$, $p = .014$, $R^2 = 0.005$, respectively). During the emotion regulation block for both groups, ERN amplitude was larger during Black-prime compared to White-prime trials: for Reappraisal, $M_{\text{Black}} = -3.66$, $SE = 0.59$, CI [-4.46, -2.14], and $M_{\text{White}} = -2.52$, $SE = 0.59$, CI [-4.81, -2.50], $t[1496] = -5.44$, $p < .001$, CI [1.54, 0.72]; for Suppression, $M_{\text{Black}} = -1.96$, $SE = 0.69$, CI [-3.31, -0.62], and $M_{\text{White}} = -1.17$, $SE = 0.69$, CI [-2.51, 0.17], $t(1347) = 3.76$, $p < .001$, CI [1.21, 0.38]. This is in contrast to the Attend block, during which ERN amplitudes were similar for both Black-prime and White-prime trials in both groups of participants. Specifically, for participants who completed Reappraisal as their first block, ERN amplitudes during the Attend (second) block differed marginally by prime race, $M_{\text{Black}} = -2.93$, $SE = 0.59$, CI [-4.09, -1.77]; $M_{\text{White}} = -3.30$, $SE = 0.59$, CI [-3.68, -1.38]; $t(1496) = 1.75$, $p = .081$, and for participants who completed Suppression as their first block, ERN amplitudes during the Attend (second) block did not differ at all by prime race, $M_{\text{Black}} = -1.73$, $SE = 0.69$, CI [-3.09, -0.38]; $M_{\text{White}} = -1.73$, $SE = 0.69$, CI [-3.08, -0.37]; $t(1347) = 0.03$, $p = .976$. No other effects were significant in the model examining ERN amplitudes in the Reappraisal group, *F*s < 2.5, *p*s > .10.

The model for the Suppression group yielded an additional significant effect of target, $F(1, 1347) = 73.87$, $p < .001$, $R^2 = 0.052$, and a marginal Prime × Target interaction, $F(1, 1347) = 3.40$, $p = .066$, $R^2 = 0.003$. These effects were qualified by a significant Block × Prime × Target interaction, $F(1, 1347) = 17.86$, $p < .001$, $R^2 = 0.013$. To probe this three-way interaction, we tested the Prime × Target interaction separately within the Suppression and Attend blocks. The Prime × Target interaction for the Suppression block was significant, $F(1, 639) = 21.53$, $p < .001$, $R^2 = .033$. ERNs during Black-tool, Black-gun and White-tool errors did not differ from one another (*t*s < 1, *p*s > .70; *M*s range from -2.28 to -2.13, *SE*s range from 0.75 to 0.76, CIs fall within -3.77 to -0.65), whereas White-gun errors elicited a smaller ERN ($M = -0.26$, $SE = 0.76$, CI [-1.74, 1.23]), which differed from both White-tool, $t(639) = 6.66$, $p < .001$, CI [1.42, 2.62], and Black-gun trials, $t(639) = 6.23$, $p < .001$, CI [1.29, 2.47]. The Prime × Target interaction during the Attend block was marginal, $F(1, 486) = 3.42$, $p = .065$, $R^2 = .007$, but reflected a more typical pattern for the WIT: the ERN during Black-tool errors was greater than that during Black-gun errors ($\Delta = -1.90$, CI [-2.49, -1.31]), $t(486) = -6.36$, $p < .001$, but ERNs during White-tool and White-gun errors did not differ ($\Delta = -0.45$, CI [-1.04, 0.15]), $t[486] = -1.48$, $p = .139$. This model also showed a significant effect of target, $F(1, 486) = 48.26$, $p < .001$, $R^2 = .090$, resulting from larger ERN amplitudes during tool ($M = -2.51$, $SE = 0.78$, CI [-4.03, -0.97]) than gun trials ($M = -1.01$, $SE = 0.78$, CI [-2.54, 0.52]).

We also examined the Block × Target effect separately at each level of prime. We found significant two-way interactions within both Black-prime ($F[1, 558] = 16.35$, $p < .001$, $R^2 = 0.028$) and White-prime trials ($F[1, 531] = 5.97$, $p < .001$, $R^2 = 0.011$), both generated by distinct Block × Target patterns. Most relevant to our hypotheses was the Block × Prime interaction within the Black-prime trials, which reflected that ERN amplitude for Black-gun trials decreased from the Suppression block ($M = -1.68$, $SE = 0.82$, CI [-3.29, -0.06]) to the Attend block ($M = -0.69$, $SE = 0.84$, CI [-2.33, 0.96]), $t(558) = 2.84$, $p = .005$, CI [0.31, 1.68], while ERN amplitude increased across those blocks for Black-tool trials (*M*s = -1.89 and -2.75, *SE*s = 0.82 and 0.83, CIs [-3.50, -0.28] and [-4.38, -1.13], respectively), $t(558) = 2.80$, $p = .005$, CI [0.26, 0.86]. The latter result is most relevant to our hypotheses, suggesting that the distress response to race bias related errors (i.e., Black-tool) was greater during the Attend block compared to the Suppression block. However, because block instructions were perfectly confounded with block order, it is not possible from these analyses to separately examine the effect of emotion regulation strategies.

Testing block order effects

Between-subjects comparisons using data from the first blocks of both experiments revealed a number of differences from the (within-subjects) findings of Experiment 1. In particular, whereas completing the emotion-regulation blocks after the Attend block (Experiment 1) led to increased bias and reduced PDP-C and ERN amplitude, none of those effects emerged when the emotion-regulation blocks were completed first. These differences suggest that order effects, and possibly their interaction with instructions, affect control-related processes in the WIT. This could occur because of fatigue, in which case response accuracy (and possibly ERN amplitude) should be lower in the second trial block regardless of which instructions are given. Alternatively, attempting to implement an emotion regulation strategy might make the task more challenging, in which case accuracy (and, perhaps, ERN amplitude) should be lower in the second block of trials only when suppression or reappraisal are attempted in that block.

These possibilities were tested using a set of 2 (Block; block 1, block 2) × 3 (Instructions; Attend, Reappraisal, Suppression) mixed models with random intercepts grouped by subject (and, for the ERN, electrodes within subjects). The model of overall WIT accuracy showed a main effect of block, $F(1, 105) = 22.63$, $p < .001$, $R^2 = 0.18$, indicating that accuracy was greater during the first ($M = .74$, $SE = .01$, CI [.71, .77]) compared to the second block ($M = .69$, $SE = .01$, CI [.66, .72]). The main effect of instructions was not significant ($F < 1$, $p > .6$); however, there was a significant Instructions × Block interaction, $F(2, 105) = 3.62$, $p = .030$, $R^2 = .065$. Simple effect tests showed that whereas accuracy under Attend instructions was similar whether Attend occurred as block 1 or block 2 (*M*s = .71 and .72, *SE*s = .02 and .02, respectively; $t < 1$, $p > .8$, CI [-.05, .04]), accuracy under Suppression instructions was greater when Suppression occurred during the first block ($M = .76$, $SE = .02$) compared to the second block ($M = .66$, $SE = .02$), $t(105) = 3.73$, $p < .001$, CI [.05, .16]. A similar trend was evident for Reappraisal (*M*s = .74 and .69, *SE*s = .02 and .02, respectively), though the difference was marginal, $t(105) = 1.81$, $p = .073$, CI [.00, .11].

Finally, the model examining ERN amplitudes showed a significant main effect of block, $F(1, 5499) = 6.96$, $p = .008$, $R^2 = 0.001$, indicating that ERNs measured during the first block ($M = -2.09$ µV, $SE = 0.31$, CI [-2.69, -1.49]) were larger than those during the second block ($M = -1.52$ µV, $SE = 0.33$, CI [-2.17, -0.87]), $\Delta = -0.57$, CI [-1.00, -0.15]. There was also a significant main effect of instructions, $F(2, 5499) = 9.09$, $p < .001$, $R^2 = 0.004$. Overall, the magnitude of ERNs during Attend ($M = -2.04$ µV, $SE = 0.30$, CI [-2.64, -1.45]) were greater than during both Reappraisal ($M = -1.74$ µV, $SE = .31$, CI [-2.36, -1.13]), $t(5499) = -2.61$, $p = .010$, CI [-0.52, -0.07], and Suppression ($M = -1.62$ µV, $SE = 0.32$, CI [-2.24, -1.01]), $t(5499) = -3.40$, $p < .001$, CI [-0.66, -0.18]. The Instructions × Block interaction was not significant, $F(2, 5499) = 1.93$, $p = .145$, $R^2 < 0.001$.

These analyses suggest that accuracy was reduced in the second block of trials compared to the first, but only if that second block involved emotion regulation. Interestingly, completing an emotion regulation block first had no effect on performance in a subsequent Attend block (compared to when Attend was completed first), suggesting that simple fatigue does not account for block order effects. These analyses also provide a different perspective on the predicted reduction of the ERN when engaging in emotion regulation. When collapsed across block order, the significant main effect of instructions in the ERN model supports our hypothesis that emotion regulation strategies would reduce the ERN. This is consistent with findings reported by Hobson et al. (2014), whose analyses also collapsed across block order.

General discussion

Recent theorizing has characterized the ERN as a reflection of a neural distress signal, emanating from dACC and other salience network structures, engaged when behavior falls short of task goals (see Eisenberger & Lieberman, 2004; Inzlicht et al., 2015). In situations where errors can signify the expression of race bias (Amodio et al., 2004, 2008) this neural distress signal could be conceptually related to the compunction felt by well-meaning individuals when their behavior reveals bias (Devine et al., 1991; Monteith, 1993; Monteith et al., 2002, 2010). Alcohol ingestion, long known to reduce anxiety in distressing situations (see Greeley & Oei, 1999), reduces ERN amplitude and disrupts control-related behavior (Bartholow et al., 2012; Ridderinkhof et al., 2002), leading to increased response bias in the WIT (Bartholow et al., 2012; Schlauch, Lang, Plant, Christensen, & Donohue, 2009). The purpose of the current work was to investigate whether conscious implementation of expressive suppression or cognitive reappraisal might have similar effects.

Results of the first experiment, in which emotion regulation was manipulated within subjects, appeared to support this possibility. Compared to the control (Attend) condition, during the emotion regulation blocks accuracy bias was more pronounced and ERN amplitudes were smaller overall but were larger for Black- relative to White-primed trials. Moreover, the pattern of findings was consistent with the prediction that Reappraisal would be more effective than Suppression at reducing error-related distress (Gross, 2002; Ochsner et al., 2012), and thus would produce more biased responding compared to Suppression. Specifically, under Reappraisal instructions participants expressed more race bias in their behavior and their errors elicited smaller ERNs compared to both Attend and Suppression conditions. Reappraisal also was associated with reduced PDP-C relative to Attend but not relative to Suppression, indicating partial support for this prediction. Finally, in comparison to both other conditions, Reappraisal also led participants to rely more on automatic associations (i.e., PDP-A), consistent with the general notion of Reappraisal contributing to a more biased pattern of responding. Overall, then, although cognitive reappraisal generally is considered an adaptive strategy for managing distress, the findings from the first experiment suggested that engaging in Reappraisal can have maladaptive consequences in the context of failed attempts to control race bias.

However, the design of Experiment 1, in which participants always completed the task under Attend instructions first, complicates interpretation of the findings due to the potential for order effects to influence the data. The design of the second experiment permitted a number of comparisons to isolate effects of Suppression and Reappraisal. In the fully between-subjects comparisons utilizing data from the first blocks of both experiments, a number of effects that emerged from the within-subjects comparisons did not replicate. First, when each was completed as participants' first trial block, performing the WIT under Suppression or Reappraisal instructions did not appear to increase accuracy bias compared to the Attend instructions. There also were no effects of instructions on PDP-C or PDP-A estimates when first-block performances were compared, suggesting that emotion-regulation instructions neither reduced controlled processing nor enhanced reliance on automatic associations. The between-subjects behavioral results suggesting no increase in WIT errors under emotion regulation conditions is consistent with findings from Ichikawa et al. (2011) and Hobson et al. (2014), who also found no effect of emotion regulation instructions on response accuracy.

Unlike Experiment 1, the fully between-subjects ERN analyses suggested that while Suppression failed to reduce the ERN signal overall, there was a reduction in the ERN specifically for bias-related errors relative to other error types. This may suggest that these participants were less bothered by expressing bias in particular. In the between-subjects analysis, we found that Reappraisal actually increased ERN amplitude during Black-prime trials specifically, which again is inconsistent with our predictions and the within-subjects results from Experiment 1.

These differences in the between- and within-subjects results raise the possibility that the deleterious effects of emotion regulation seen in the first experiment arose because of factors related to completing the WIT more than once (e.g., fatigue), which could decrease the ability to implement control (Lorist, Boksem, & Ridderinkhof, 2005). However, models testing this possibility showed that simply having completed the task already did not lead to changes in overall performance. Accuracy was similar under Attend instructions regardless of whether the Attend block was completed first or following an emotion-regulation block. Rather, accuracy was reduced during the second block only when that block required emotion regulation, particularly Suppression. This suggests a more nuanced picture of the block order effects in the data, whereby second-block performance suffers as a result of the task becoming more challenging due to the requirement to implement Suppression (or, to a lesser extent, Reappraisal) after having already completed the task without such a requirement. Conversely, completion of the Attend block following an emotion regulation block should be less challenging because participants are no longer required to regulate.

Unlike these differences in patterns of behavioral responses across experiments, some effects of emotion regulation instructions on patterns of ERN response appeared robust to block order. Across both experiments, regardless of whether it occurred first or after an Attend block, engaging in Suppression or Reappraisal led to larger ERN amplitudes for errors occurring on Black- relative to White-primed trials. No such prime race effects emerged during Attend blocks. That this effect occurred only for emotion-regulation conditions is consistent with previous findings showing that emotion regulation enhances error-related neural activity (Bratec et al., 2015). These patterns occurred regardless of the overall amplitude of the ERN, which was larger during participants' first blocks compared to their subsequent blocks (see Supplementary Material). That is, despite potential fatigue from performing the WIT more than once, Reappraisal and Suppression were associated with differentially large ERNs on Black-primed trials. One possible explanation for this effect is that attempting to reappraise or suppress affective reactions to errors is less difficult when errors have less obvious implications for race bias (i.e., on White-prime trials). That is, mistaking a tool for a gun following a White face likely elicits less compunction than a similar error following a Black face (Monteith, 1993; Monteith et al., 2002, 2010), making reappraising or suppressing that feeling easier to manage. Alternatively, larger ERNs on Black relative to White prime trials during the emotion-regulation blocks could reflect increased monitoring on trials when a cue (i.e., Black prime) signals that an error is more likely, or that an error will require increased effort to regulate emotion (van Noordt, Desjardins, & Segalowitz, 2015). This could have occurred because, unlike in some previous studies (e.g., Hobson et al., 2014), the instructions in the two emotion-regulation blocks focused on regulating responses to errors in particular (as opposed to asking participants to "adopt a detached as you complete the task").

The current findings have a number of implications for research aimed at understanding effects of emotion regulation on behavioral and brain responses. The between-subjects comparisons reported here raise questions concerning prior reports from within-subjects comparisons indicating that emotion-regulation instructions reduced error-related neural responses (Hobson et al., 2014; Ichikawa et al., 2011). Here, there was no evidence of reduced ERN under Suppression or Reappraisal when those instructions were implemented during participants' first trial block. Although Ichikawa et al. (2011) reported that medial ACC regions involved in error processing can be modulated by participant-chosen emotion regulation strategies, most of the error-related activation differences they reported occurred too late after the response (~ 9 s) to have been related to the ERN. Hobson et al. (2014) found that a detachment reappraisal strategy was effective in reducing ERN amplitude relative to the no-instruction control condition. It is possible that this detached strategy is more effective at limiting error-related distress compared to a cognitive reappraisal strategy, which requires focusing on errors in order to reappraise them. Such attention to errors could ironically enhance error-related distress in the short-term – i.e., the immediate reaction reflected in the ERN – even if such a strategy could reduce the consequences of that distress later – even just 9 s later (Ichikawa et al., 2011).

It is also important to bear in mind that, unlike in previous studies (Hobson et al., 2014; Ichikawa et al., 2011), the task used in the current studies produces errors indicative of *racial bias*. Errors committed during the WIT and similar tasks are highly salient because of their potential to reveal biased attitudes that conflict with most participants' stated beliefs (see Amodio et al., 2004). In other words, although committing errors is aversive regardless of the context (Hajcak & Foti, 2008), the conflict associated with errors indicative of underlying racial biases might be particularly acute and difficult to regulate by means of Reappraisal.

Conclusions

The present results suggest that it may be more difficult to modulate neurophysiological responses to errors using emotion-regulation instructions than suggested in previous research (Hobson et al., 2014; Ichikawa et al., 2011). This might be particularly true for fast-paced tasks such as the WIT, where performance accuracy has implications for sensitive racial issues. Researchers should continue to investigate the affective properties of error monitoring, including whether attempts at affect regulation differentially modulate responses elicited by different types of errors. Additionally, the importance of the block order effect in determining neural and behavioral responses in these studies suggests that effects of emotion regulation might best be tested using between-subjects designs. Related to that point, future research should seek to further investigate how mental fatigue and/or task order impact performance during emotion regulation tasks.

Author note Support for this research was provided in part by grant R01 AA020970 from the National Institute on Alcohol Abuse and Alcoholism. The authors declare no conflicts of interest related to this research.

References

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: Evidence from ERPs. *Emotion*, *13*, 960-973.

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*, 60-74.

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*, 88-93.

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience, 1*, 26-36.

Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology, 37*, 13–20.

Bartholow, B. D., Henry, E. A., Lust, S. A., Saults, J. S., & Wood, P. K. (2012). Alcohol effects on performance monitoring and adjustment: affect modulation and impairment of evaluative cognitive control. *Journal of Abnormal Psychology*, *121*, 173-186.

Bebko, G. M., Franconeri, S. L., Ochsner, K. N., & Chiao, J. Y. (2011). Look before you regulate: differential perceptual strategies underlying expressive suppression and cognitive reappraisal. *Emotion*, *11*, 732.

Braem, S., King, J.A., Korb, F.M., Krebs, R.M., Notebaert, W., & Egner, T. (2017). The role of anterior cingulate cortex in the affective evaluation of conflict. *Journal of Cognitive Neuroscience, 29*(1), 137-149.

Bratec, S. M., Xie, X., Schmid, G., Doll, A., Schilbach, L., Zimmer, C., ... & Sorg, C. (2015). Cognitive emotion regulation enhances aversive prediction error activity while reducing emotional responses. *Neuroimage*, *123*, 138-148.

Cavanagh, J.F. & Shackman, A.J. (2015) Frontal midline theta reflects anxiety and cognitive control: Meta analytic evidence. *Journal of Physiology – Paris, 109,* 3-15.

Carter, C.S., & van Veen, V. (2007) Anterior cingulate and conflict detection: An update of theory and data. *Cognitive, Affective and Behavioral Neuroscience, 7*, 367-379.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality and Social Psychology*, *60*, 817-830.

Dreisbach, G., & Fischer, R. (2015). Conflicts as aversive signals for control adaptation. *Current Directions in Psychological Science*, *24*(4), 255-260.

Edwards, L. J., Muller, K.E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 Statistic for fixed effects in the linear mixed model. *Statistics in Medicine, 27*, 6137-6157.

Eisenberger, N. I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience, 13*, 421-434. doi:https://doi.org/10.1038/nrn3231

Goldin, P. R., McRae, K., Ramel, W., & Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, *63*, 577-586.

Greeley, J., & Oei, T. (1999). Alcohol and tension reduction. In K. E. Leonard & H. T. Blane (Eds.), Psychological theories of drinking and alcoholism (2nd ed., pp. 14 –53). New York: Guilford Press.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*, 271-299.

Gross, J. J. (2001). Emotion regulation in adulthood: Timing is everything. *Current Directions in Psychological Science*, *10*, 214-219.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*, 281-291.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*, 348-362.

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological Science*, *19*, 103-108.

Ham, T., Leff, A., de Boissezon, X., Joffe, A., & Sharp, D. J. (2013). Cognitive control and the salience network: An investigation of error processing and effective connectivity. *The Journal of Neuroscience, 33*, 7091-7098.

Hobson, N. M., Saunders, B., Al-Khindi, T., & Inzlicht, M. (2014). Emotion down-regulation diminishes cognitive control: A neurophysiological investigation. *Emotion*, *14*(6), 1014.

Ichikawa, N., Siegle, G.J., Jones, N.P., Kamishima, K., Thompson, W.K., Gross, J.J., & Ohira, H. (2011). Feeling bad about screwing up: emotion regulation and action monitoring in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience, 11*, 354-371.

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in cognitive sciences*, *19*(3), 126-132.

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

Lieberman, M. D., & Eisenberger, N. I. (2015). The dorsal anterior cingulate cortex is selective for pain: Results from large-scale reverse inference. *Proceedings of the National Academy of Sciences*, *112*(49), 15250-15255.

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research*, *24*, 199-205.

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J., & Gross, J. J. (2008). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group Processes & Intergroup Relations*, *11*(2), 143-162.

Monteith, M. J. (1993). Self-regulation of prejudiced responses: Implications for progress in prejudice-reduction efforts. *Journal of Personality and Social Psychology*, *65*, 469-485.

Monteith, M. J., Ashburn-Nardo, L., Voils, C. I., & Czopp, A. M. (2002). Putting the brakes on prejudice: On the development and operation of cues for control. *Journal of Personality and Social Psychology, 83*, 1029-1050.

Monteith, M.J., Mark, A.Y., & Ashburn-Nardo, L. (2010). The self-regulation of prejudice: Toward understanding its lived character. *Group Processes and Intergroup Relations, 13*, 183-200.

Ochsner, K. N., Bunge, S. A., Gross, J. J., & Gabrieli, J. D. (2002). Rethinking feelings: An fMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, *14*, 1215-1229.

Ochsner, K. N., Silvers, J. A., & Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, *1251*, E1-E24.

Opitz, P. C., Rauch, L. C., Terry, D. P., & Urry, H. L. (2012). Prefrontal mediation of age differences in cognitive reappraisal. *Neurobiology of Aging*, *33*(4), 645-655.

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.

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.

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon identifications. *Journal of Experimental Social Psychology, 41*, 36-47.

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science*, *298*(5601), 2209-2211.

Schlauch, R. C., Lang, A. R., Plant, E., Christensen, R., & Donohue, K. F. (2009). Effect of alcohol on race-biased responding: The moderating role of internal and external motivations to respond without prejudice. *Journal of Studies on Alcohol and Drugs, 70*, 328–336.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P3 ERP. *Psychophysiology*, *23*, 695-703.

Spunt, R. P., Lieberman, M. D., Cohen, J. R., & Eisenberger, N. I. (2012). The phenomenology of error processing: the dorsal ACC response to stop-signal errors tracks reports of negative affect. *Journal of Cognitive Neuroscience*, *24*, 1753-1765.

Tibon, R., & Levy, D. A. (2015). Striking a balance: Analyzing unbalanced event-related potential data. *Frontiers in Psychology, 6*.

van Noordt, S. J., Desjardins, J. A., & Segalowitz, S. J. (2015). Watch out! Medial frontal cortex is activated by cues signaling potential changes in response demands. *NeuroImage*, *114*, 356-370.

van't Wout, M., Chang, L. J., & Sanfey, A. G. (2010). The influence of emotion regulation on social interactive decision-making. *Emotion*, *10*(6), 815.

Volpert-Esmond, H.I., Merkle, E.C., Levsen, M.P., Ito, T.A., & Bartholow, B.D. (2018). Using trial-level data and multilevel modeling to investigate within-task change in event-related potentials. *Psychophysiology, 55*, e13044.

Vossen, H., Van Breukelen, G., Hermens, H., Van Os, J., & Lousberg, R. (2011). More potential in statistical analyses of eventrelated potentials: A mixed regression approach. *International Journal of Methods in Psychiatric Research, 20*(3), e56–e68.

West, B. T., Welch, K. B., & Galecki, A. T. (2015). Linear mixed models: A practical guide using statistical software (2nd ed.). Boca Raton, FL: CRC Press.