



# Proactive engagement of cognitive control modulates implicit approach-avoidance bias

Katia M. Harlé<sup>1,2</sup> · Jessica Bomyea<sup>1,2</sup> · Andrea D. Spadoni<sup>1,2</sup> · Alan N. Simmons<sup>1,2</sup> · Charles T. Taylor<sup>2</sup>

Published online: 5 August 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

Implicit social-affective biases—reflected in a propensity to approach positive and avoid negative stimuli—have been documented in humans with paradigms, such as the Approach-Avoidance Task (AAT). However, the degree to which preemptively engaging cognitive control can help to down-regulate those behavioral tendencies remains poorly understood. While undergoing functional magnetic resonance imaging (fMRI), 24 healthy participants completed a cued version of the AAT, in which they responded to pictures of happy or angry faces by pulling a joystick toward themselves (approach) or pushing the joystick away (avoidance) based on the color of the stimulus frame. On some trials, they were cued to reverse the frame color/joystick action instructions. Before stimulus onset, a reverse cue was associated with deactivation of a visuo-spatial and motor planning network and subsequent slowing down in response to stimuli. During the stimulus phase, a reverse cue was associated with a) activation of cognitive control areas, including the right inferior frontal gyrus (IFG) and right inferior parietal lobule (IPL); and b) reduced right precentral gyrus activation when having to push (avoid) a happy face. Overall, these results suggest that proactively engaging cognitive control can help fine-tune behavioral and neural adjustment to emotionally incongruent behavioral conditions.

**Keywords** Approach-avoidance task · Cognitive control · Proactive control · AAT · CAAT

## Introduction

Humans exhibit automatic behavioral tendencies to approach or avoid relevant social stimuli. For example, an intrinsic tendency to approach positive social stimuli (e.g., smiling face) and to avoid negative cues (e.g., angry person) has been well documented in implicit bias tasks, such as the approach-avoidance task (AAT) (Chen & Bargh, 1999; Rinck & Becker, 2007; Roelofs, Elzinga, & Rotteveel, 2005). Specifically, healthy, euthymic individuals tend to be faster at executing these affect-congruent actions, while they are slower in incongruent situations (i.e., approaching a negative cue or avoiding a positive one). Such valence-action

tendencies are consistent with the influence of hard-wired, Pavlovian biases on instrumental learning of such actions (Chiu, Cools, & Aron, 2014; Crockett, Clark, & Robbins, 2009; Guitart-Masip et al., 2011; Guitart-Masip et al., 2012). For instance, individuals learn more successfully to approach a stimulus when anticipating a reward rather than the absence of punishment, and they learn better to withhold action when anticipating a punishment rather than the absence of reward (Guitart-Masip et al., 2012). These behavioral biases are associated with anticipatory up- or down-regulation of motor excitability in response to appetitive or aversive cues, respectively (Chiu et al., 2014). Neurocognitive modeling of such biases therefore may not only explain general abnormalities in approach-avoidance behavior, it may help to deepen our understanding of how these biases may be modified with targeted neurocognitive interventions, e.g., in individuals with maladaptive approach-avoidance tendencies.

In affective bias paradigms, such as the AAT, the approach is usually operationalized as pulling cues toward self (e.g., with a joystick), thereby bringing them closer, whereas avoidance is operationalized as pushing away the stimulus, which becomes more distant. While behavioral incongruency effects (e.g., in terms of motor response times) are not always reliably observed (Derntl et al., 2011; Ernst et al., 2013; Korucuoglu,

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.3758/s13415-020-00815-3>) contains supplementary material, which is available to authorized users.

✉ Katia M. Harlé  
kharle@ucsd.edu

<sup>1</sup> VA San Diego Healthcare System, 3350 La Jolla Village Drive, San Diego, CA 92161, USA

<sup>2</sup> Department of Psychiatry, University of California San Diego, La Jolla, CA, USA

Gladwin, & Wiers, 2014; Phaf, Mohr, Rotteveel, & Wicherts, 2014; Seidel et al., 2010), neural investigations of the AAT have identified key regions involved in the processing of incongruent contexts in the task (i.e., having to approach a negative social stimulus or to avoid a positive stimulus). Specifically, the ventro-lateral prefrontal cortex (PFC), including the inferior frontal gyrus (IFG), and the dorso-lateral PFC (DLPFC) are recruited to summon cognitive control and override behavioral tendencies in response to incongruency (Bertsch et al., 2018; Derntl et al., 2011; Ernst et al., 2013; Roelofs, Minelli, Mars, van Peer, & Toni, 2009; Volman, Roelofs, Koch, Verhagen, & Toni, 2011a; Volman, Toni, Verhagen, & Roelofs, 2011b). These studies highlight a complex regulatory process, which includes working memory (supported by recruitment of the DLPFC) and inhibitory control, i.e., response or attentional inhibition (as reflected by IFG recruitment); both processes are subsumed within the broader construct of cognitive control (Miyake et al., 2000).

Not surprisingly, this type of paradigm has been used effectively to identify subtle alterations in affective processing bias among individuals with a range of psychopathology (Loijen, Vrijnsen, Egger, Becker, & Rinck, 2020). For instance, anxious individuals are slower at pulling fear-triggering stimuli toward themselves—analogue to behavioral approach—relative to neutral pictures (Heuer, Rinck, & Becker, 2007; Najmi, Kuckertz, & Amir, 2010; Rinck & Becker, 2007), while depression symptoms are associated with a faster tendency to push away negative stimuli (Derntl et al., 2011; Eisma et al., 2015; Seidel et al., 2010). Overall, this research highlights the potential clinical utility of implicit affective bias tasks, such as the AAT, to identify predictive characteristics of psychopathology and treatment targets for cognitive training and exposure therapy, e.g., via repeated practice of incongruent associations (Jacobus et al., 2018; Rinck, Wiers, Becker, & Lindenmeyer, 2018; Taylor & Amir, 2012; Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011; Wiers, Rinck, Kordts, Houben, & Strack, 2010). Some studies, however, have failed to find any beneficial effects of approach-avoidance training (Becker, Jostmann, Wiers, & Holland, 2015; Loijen et al., 2020; van Uijen, van den Hout, & Engelhard, 2015), suggesting that a deeper understanding of the processes that underpin implicit affective biases is needed. To bridge this gap, the present study seeks to identify reliable neurocognitive training targets for implicit socio-affective biases, starting in a relatively healthy normative sample, and with the long-term goal of applying such paradigms to clinical populations with more pronounced biases.

So far, investigations of the AAT in both healthy and clinical populations have focused on the reactive processing and regulation of social affective cues and do not address whether engaging proactive cognitive control can modulate prepotent reactive behavioral tendencies. From a dual mechanisms framework of cognitive control (Braver, 2012), reactive control is a transient, late corrective mechanism, which starts at

the onset of conflicting stimulus presentation to resolve interference. In contrast, proactive control involves the anticipation and prevention of interference prior to being presented with the conflicting situation. Understanding the role of such proactive, top-down modulation of cognitive control on more reactive socio-affective biases therefore has important implications for assessing and potentially training top-down regulation of these automatic tendencies. For instance, proactive control may be more amenable to introducing, manipulating, and maintaining in memory new rules and expectations, which may in turn help to compensate for reactive cognitive control biases, such as processing of affectively incongruent responses in the AAT. While both types of cognitive control processes may be engaged independently or simultaneously, biases toward one or the other may vary across subclinical and psychiatric conditions (Braver, 2012; Braver, Paxton, Locke, & Barch, 2009). Recognizing such differences may be particularly useful in psychiatric populations to help develop and enhance cognitive bias training protocols to help individuals change maladaptive social-affective biases (Brevers et al., 2018; Fales, Barch, Rundle, et al., 2008b; Lesh et al., 2013; West, Choi, & Travers, 2010).

In terms of negative affective pathology, impairments in reactive processing and resolution of emotional conflict have been more consistently observed in depression (Fales, Barch, Rundle, et al., 2008b; Holmes & Pizzagalli, 2007; Saunders & Jentsch, 2014) and anxiety (Fales, Barch, Burgess, et al., 2008a; Krug & Carter, 2010). There have been mixed findings in terms of proactive control in negative affective pathology, with some studies showing deficits and reduced prefrontal recruitment in anxiety (Ansari & Derakshan, 2011; Fales, Barch, Burgess, et al., 2008a), but others showing no impairments (Ng, Chan, & Schlaghecken, 2012). While reactive control processes, by nature more implicit and automatic, may be more difficult to modulate with training, depressed or anxious individuals may still be successfully trained to recruit proactive control and thus benefit from such top-down, goal-directed form of cognitive control recruitment. In other words, some individuals may still be able to use proactive cognitive control to help compensate for the more consistently observed reactive control abnormalities in this population.

Given this body of research and the outlined relevance of the proactive versus reactive control distinction to implicit affective bias tasks, the goal of this study was to: 1) examine the role of cued/proactive cognitive control on approach and avoidance action tendencies in a social affective context; and 2) understand the neural pathway of such mechanisms. Specifically, we aimed to disentangle the neural impact of proactive vs reactive cognitive control on individuals' processing of affective incongruence in the AAT. As in other interference tasks probing the flexibility of cognitive control (e.g., with switching instruction sets) (Delis, Kaplan, & Kramer, 2001), we assessed the effect of proactively preparing

to reverse instructions associated with the more automatic, prepotent stimulus-response set, rather than cueing for the specific types of stimulus-response instruction to be encountered and regulated.

We hypothesized that engaging cognitive control proactively by cueing to reverse typical stimulus/response rules in the AAT would result in 1) a more careful stimulus-response implementation, i.e., slower reaction times, and 2) stronger engagement of inhibitory control areas, particularly the IFG, in response to incongruent stimulus-action trials. Given the research highlighted above linking negative affective pathology and deficits in reactive cognitive control and affective conflict processing, determining whether negative affect modulates proactive cognitive control recruitment in the AAT may inform the development and improvement of cognitive behavioral treatments for psychiatric disorders. Thus, a third exploratory goal of this study was to evaluate the role of negative affect in these processes, i.e., whether individuals with elevated negative affect may be more or less amenable to cued modulation of the neural systems supporting effective, flexible cognitive control.

## Methods

### Participants

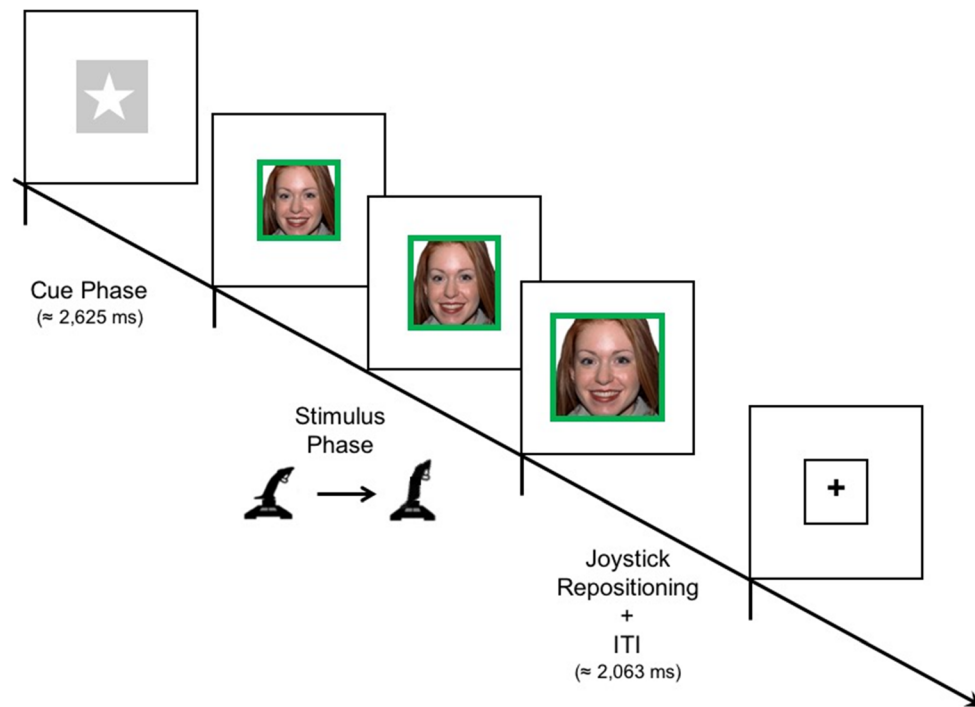
Twenty-four individuals (age range: 18–26 years, 63% women) participated in this study (approved by the Human Research Protections Program at University of California, San Diego, CA). Mean age was 22 (standard deviation [SD] = 2.8) years, and they had an average of 15.8 (SD = 1.5) years of education. Five (21%) participants were Hispanic, while the rest of the sample was non-Hispanic. Participants were recruited through flyers posted around the UC San Diego campus. All participants signed informed consent and were compensated \$50 for completing the study. Before participants performed the experimental task, they were administered the Positive and Negative Affect Scale (PANAS) (Watson, Clark, & Tellegen, 1988). Based on this measure, participants' trait positive affect ranged from 17 to 46 (mean = 33.4, SD = 6.7) and their trait negative affect ranged from 10 to 23 (mean = 14.4, SD = 3.5). Participants completed other functional MRI tasks as part of a larger test battery (data published elsewhere (Bomyea, Taylor, Spadoni, & Simmons, 2018)).

### Cued approach avoidance task

The Cued Approach Avoidance Task (CAAT) was designed based on the AAT task (see Najmi et al., 2010; Rinck & Becker, 2007; Taylor & Amir, 2012) in which participants are presented with a series of pictures on a computer screen, while a joystick was situated in their dominant hand. On each

trial, participants were presented with a happy or angry face picture framed by either a green or a blue border, instructing them to either pull the joystick if the border was green, or push the joystick if the border was blue. Thus, participants were asked to respond only to the frame color rather than to the image itself. Face pictures were selected from the Radboud database with both happy and angry emotions nested within each model (Langner et al., 2010). All faces were of white ethnicity and there were equal numbers of male and female faces. Pulling the joystick was associated with the face stimulus picture becoming increasingly larger to simulate an approach action. In contrast, pushing the joystick was associated with the face stimulus becoming increasingly smaller to simulate withdrawal away from the stimulus (avoidance). This changing frame of reference (i.e., face getting bigger or smaller) when manipulating the joystick is a key feature to simulate approach and avoidance actions for the participant. Flexor movement (pulling forearm toward oneself) and extensor movement (pushing forearm away) have been respectively associated with avoidance and approach in standard go/no-go tasks (Chiu et al., 2014; Guitart-Masip et al., 2012; Swart et al., 2018). However, there is substantial evidence that the visual feedback of one's position in relation to the stimulus when moving a joystick toward or away from oneself (i.e., seeing stimulus size increase or decrease), maps on to approach and avoidance, respectively (Cacioppo, Priester, & Berntson, 1993; Maxwell & Davidson, 2007; Phaf et al., 2014). The trial setup for the CAAT was the same as for the AAT, with the addition that each stimulus was preceded by presentation of a brief, jittered (range: 2,000 ms to 3,250 ms) visual cue, i.e., white or black star, indicating whether participants should maintain or reverse instructions to be given in the upcoming stimulus presentation. Specifically, participants were instructed as follows: "Occasionally, however, you will see a black star. When this occurs, you are required to reverse the instructions, that is, reverse the direction of the joystick according to the task instructions. So, when you see a black star, you will pull the joystick when the border is blue and push the joystick if the border is green" Thus, while participants were not primed on the specific joystick actions, they were occasionally cued to reverse instructions in order to proactively engage cognitive control with regards to task rules. The subsequent stimulus-bound phase was used to assess reactive control processes, as typically captured in the AAT to gauge implicit affective biases. In this case, we assessed reactive control both independently (following "same" cue) or modulated by (following "reverse" cue) proactive control recruitment. Thus, proactive inhibition of task instructions during the cue phase and type of conflict/incongruence encountered during the reactive control phase were temporally and parametrically unconfounded (see Fig. 1 for trial timeline).

Participants first completed 12 practice trials of a single cue CAAT, in which only the white star cue was used and



**Fig. 1** CAAT trial timeline. Each trial started with presentation of the cue (jittered duration: 2,000–3,250 ms) to indicate if one should use the same or reverse instructions in the upcoming trial. In this case, a white star indicates “same” instruction. Stimulus phase: the green border indicates joystick should be pulled (approach trial). Stimulus presentation lasts until image has reached maximum (or minimum) size as one pushes or pulls as far as possible with the joystick. For each trial, the picture disappeared when the joystick reached approximately a 30° position in

either direction, regardless of whether the participant responded correctly and trial ended. Joystick repositioning phase: An empty box then appeared in the middle of the screen and joystick position was indicated by a cross on the screen. Participants had to move the joystick (cross) back in the center of a square in the middle of the screen. Duration of this phase included the time to reposition the joystick and an interstimulus interval (ITI) (total jittered duration: 1,750–2,375 ms), after which the next trial started

participants were instructed to simply pay attention to it. They then completed a set of 12 practice trials of the experimental dual-cue CAAT (white and black stars) in which they had to keep or reverse the subsequent instructions. Finally, they performed the full CAAT protocol in the scanner, completing a total 96 trials, crossed between 6 conditions: 2 Picture Valence Type (happy, angry)  $\times$  2 Border Color (green, blue)  $\times$  2 Cue Type (Reverse, Same). Trial order was randomized across participants. Half the pictures with green borders were happy faces and half were angry faces; similarly, half the pictures with blue borders were happy faces and half were angry faces. Finally, for each set of these four possible stimulus-instruction combinations, 25% of the trials were preceded by a black star (meaning “reverse instructions”) cue, indicating that participants should push the joystick if the stimulus border was green and pull the joystick if the border was blue, while 75% of trials were preceded by a white (meaning “same instructions”) cue.

## Behavioral analyses

We applied hierarchical generalized mixed-effect linear models to participants’ initial movement reaction times (RT), treating subject as a random factor and other independent variables as fixed effects (Baayen, Davidson, & Bates,

2008). Given the positive skew of the reaction times (overall and within conditions), an inverse Gaussian distribution was used to model RTs (Lo & Andrews, 2015). Initial movement RTs (time from stimulus onset to the time participants first moved the joystick outside of the central position) were the primary behavioral measure of interest as it does not include potentially confounding individual motor differences related to grasp of the joystick, and may therefore be more sensitive to cognitive control processes during action selection. Independent variables included cue type (same vs. reverse), instruction (push vs. pull), and picture valence (happy vs. angry). Inaccurate response trials were excluded from RT analyses. A similar generalized mixed-effect linear model was used to assess the effects of cue type, instruction, and valence on accuracy, assuming binomial distribution and logit link function for binary accuracy data. We report change in log likelihood ratio (which follows a chi-squared distribution) and regression coefficients (when applicable) with associated *t*-test and *p* values for any significant effect or interaction.

## fMRI analyses

**Image acquisition** Participants completed the CAAT while being scanned in a 3 Tesla General Electric (GE) scanner with

an 8-channel head array coil. Each scanning session included a three-plane scout scan, a sagittally acquired spoiled gradient recalled (SPGR) sequence to collect T1-weighted high-resolution structural images (FOV 256 cm; matrix: acquired  $192 \times 256$  matrix resampled to  $256 \times 256$ ; 172 slices; thickness: 1 mm; TR = 8 ms, TE: 3 ms, flip angle:  $12^\circ$ , inversion time = 450 ms) and one T2\*-weighted axially acquired echo-planar imaging (EPI) scan to measure blood oxygen level dependent (BOLD) signal during the task (parameters:  $3.75 \text{ mm} \times 3.75 \text{ mm} \times 3 \text{ mm}$ ;  $64 \times 64$  acquisition matrix with a 1 mm gap, TR = 1,500 ms, TE = 32 ms, flip angle =  $80^\circ$ , and 30 slices (covering the whole brain)). The length of task ranged from 579 to 702 acquisitions ( $M = 10.31 \text{ min}$ ). To complete the task, participants used a MRI-safe joystick, which they held in their hand. The computerized task was visible to them via a projected screen visible through a mirror in the head coil.

**Pre-processing:** Preprocessing, normalization to MNI coordinates, and subsequent fMRI analyses were conducted using ANTsR, a statistical interface between Advanced Normalization Tools Software (ANTs), R statistical software, and Analysis of Functional NeuroImages (AFNI) software (Cox, 1996). Preprocessing steps included removal of temporal outliers, field inhomogeneity correction, slice time correction, and temporal whitening. Motion correction and a CompCor component-based noise correction (Behzadi, Restom, Liao, & Liu, 2007) were conducted and residualized from the timeseries. Outlying acquisitions were censored from the time series (using a 2 standard deviations cutoff rule). Regressors were convolved with a canonical hemodynamic response function (HRF) and entered into a robust linear regression model (using the `lmrob` R function) to obtain normalized beta weights and associated statistics. Data were aligned to individual anatomical and MNI template.

**First-level analyses:** At the individual level, a total of ten regressors of interest were generated for specific events in the task and included in a robust timeseries regression. These included two regressors generated for the cue phase: same, reverse; and eight regressors locked in to the stimulus phase were included reflecting the 8 possible combination of conditions of interest (2 preceding cue type  $\times$  2 instruction type  $\times$  2 stimulus valence type). Errors during the stimulus phase were added as a regressor of no interest to control for any BOLD signal related to inaccurate motor response. These regressors were mean-centered and Variance Inflation Factors (VIFs) were calculated to ensure that each regressor was uniquely identifiable despite some level of potential collinearity between cue phase and stimulus regressors. Across participants, the average VIFs for each regressor ranged from 1.26 to 3.72, and maximum values ranged from 1.73 to 4.67, thus all below the recommended threshold of 5 (Ringle, Wende, & Becker, 2015). This model allowed for examination of neural activity in response to different levels of incongruency during the stimulus phase (instruction  $\times$  valence) as typically

conceptualized the in AAT, while assessing for the additional moderating effect of cue.

**Group-level analyses:** Voxel-wise coefficients for the first-level regressors were in turn entered in two mixed-effect linear models to assess each main effect and interaction contrast of interest, with subject entered as a random factor. Specifically, to investigate neural activity related to proactive, prestimulus engagement of cognitive control, a first mixed-effect linear model assessed the main effect of cue type during the cue phase. Second, to gauge reactive cognitive control neural processes in response to potential affective congruence or incongruence, a full interactive factorial model of cue  $\times$  instruction  $\times$  valence during the stimulus phase was tested. For this second model, we were specifically interested in the cue type main effect as well as any two-way or three-way interactions involving cue type to evaluate potential moderating effect of cue on subsequent stimulus processing. Based on the total number of trials per condition and our sample size, the present study was adequately powered at or above 0.80 to detect medium effect sizes for each main effect and interaction of interest in a mixed-linear model ( $d > 0.5$ ; (Westfall, Kenny, & Judd, 2014). To correct for multiple comparisons, we used a cluster threshold adjustment based on Monte Carlo simulations. Specifically, to create an optimal smoothing kernel, we first computed a three-parameter spatial autocorrelation function (with AFNI's 3dFWHMx) from each participants' detrended pre-processed data. These parameters were averaged across subject and used with AFNI's 3dClustSim function to conduct simulations. A minimum of 15 contiguous voxels was found to result in a corrected cluster-wise activation probability of  $p < 0.05$  based on a voxel-wise a priori probability of  $p < 0.001$ . For each significant cluster, average  $t$  statistics were extracted for each participant and correlated with trait positive and negative affect measures (i.e., PANAS).

## Results

### Behavioral performance

**Initial movement reaction times (RT)** Consistent with previous AAT studies, we observed a significant main effect of cue, instruction, and valence on reaction times (from stimulus onset to initial joystick movement). Specifically, participants had longer reaction times when instructed to reverse the instruction (i.e., “reverse” cue relative to “same” cue;  $t = 7.0$ ,  $p < 0.001$ ; omnibus test:  $\chi^2(1) = 49$ ,  $p < 0.001$ ), when having to push relative to pull ( $t = 4.7$ ,  $p < 0.001$ ; omnibus test:  $\chi^2(1) = 23$ ,  $p < 0.001$ ), and when presented with angry relative to happy face stimuli ( $t = 2.6$ ,  $p < 0.05$ ; omnibus test:  $\chi^2(1) = 6.8$ ,  $p < 0.01$ ; see Fig. 2A for violin plots of within-subject RT differences for each of the three main effect condition contrast).

A significant instruction X cue interaction was also observed ( $t = 3.2, p < 0.01$ ; omnibus test:  $\chi^2(1) = 9.9, p < 0.01$ ), suggesting that individuals were faster at initiating movement when having to pull than when having to push if the prestimulus cues was “same,” whereas such push/pull RT difference was not observed following a “reverse” cue (Fig. 2B). No other interaction reached statistical significance (i.e., instruction X valence, cue X valence, instruction X cue X valence;  $ps > 0.05$ ; see Table S.1. in *Supplementary Material*).

“Same” Condition: While reaction times were expected to be primarily modulated by the instruction reversal cue in the present task, we specifically assessed for any potential valence by joystick instruction congruency in trials with no reversal cue (i.e., “Same” instruction cue). Participants were found to be faster at pulling (mean RT = 695 ms, SE = 16 ms) than pushing away (mean RT = 739 ms, SE = 17 ms) happy faces ( $t = 4.7, p < 0.001$ ; omnibus test:  $\chi^2(1) = 20.8, p < 0.001$ ), consistent with higher propensity to approach than to avoid of happy faces. Moreover, while participants were generally slower at responding to angry faces ( $t = 2.9, p < 0.01$ ; omnibus test:  $\chi^2(1) = 8.2, p < 0.01$ ), they also were faster at pulling (mean RT = 707 ms, SE = 15 ms) than pushing away (mean RT = 793 ms, SE = 19 ms) angry faces ( $t = 4.6, p < 0.001$ ; omnibus test:  $\chi^2(1) = 20.9, p < 0.001$ ). Neither negative nor positive affect measures further modulated reaction times ( $ps > 0.05$ ).

**Performance accuracy** The main effect of instruction (Push vs. Pull) on trial accuracy was statistically significant, such that participants were overall less likely to be accurate on push trials (odds ratio = 0.35, Wald  $z = -3.3, p < 0.001$ ; omnibus test:  $\chi^2(1) = 12.1, p < 0.001$ ). The main effects of trial cue and valence did not reach statistical significance ( $ps > 0.05$ ). Echoing the interaction pattern observed with RTs (see above), a significant instruction X cue

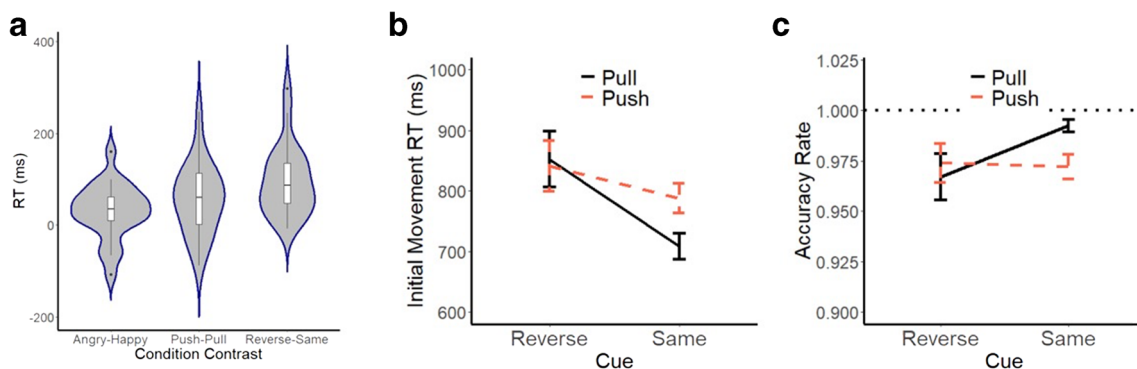
interaction was also observed (Wald  $z = 3.4, p < 0.001$ ; omnibus test:  $\chi^2(1) = 12.3, p < 0.001$ ). Specifically, individuals were more accurate when having to pull than when having to push if the prestimulus cues was “same,” but such instruction-based accuracy difference was absent following a “reverse” cue (Fig. 2C; Table S.2. in *Supplementary Material*). We note, however, that trial accuracy was generally high and limited in range (Fig. 2C), because this task is not designed to promote error but rather aims to capture reaction time effects. Negative and positive affect measures were not significantly related to accuracy ( $ps > 0.05$ ).

## fMRI analyses

### Cue phase

Six regions were identified in which a significant mean difference between reverse and same cue was observed, including the right precentral gyrus (Brodmann Area/BA 4, 95 voxels), the right cingulate gyrus/presupplementary motor area (BA 6, 53 voxels), the right (BA 22, 50 voxels) and left (BA 22, 38 voxels) superior temporal gyrus, the left lentiform nucleus (19 voxels), and the left medial prefrontal cortex (BA 6/32;  $ps < 0.001$ ; see Table 1 for more details). In all those regions, activation to “same” cue (white star presentation) was on average positive and significantly higher relative to activation to “reverse” cue (black star presentation). Thus, overall, a reduced recruitment of these regions was observed when participants were cued to reverse the upcoming instruction. No statistically significant relationship between such activation coefficients and positive or negative affect was observed.

To further assess how such differential cue activation related to motor preparedness and subsequent performance, and given the high degree of collinearity between ROI cue-related



**Fig. 2** **A.** Violin plots of the difference in initial movement reaction times (RT) for each main effect contrast (i.e., Valence: Angry-Happy faces; Border Instruction: Push-Pull; Preceding Cue: Reverse-Same). **B.** Interaction between cue and instruction on initial movement reaction times. Reaction times did not differ between push and pull trials following reverse cues, whereas reaction time was on average significantly faster on pull trial in “same” cue trials. We note that the same three main

effects of cue, instruction, and valence, as well as the instruction X cue interaction were statistically significant and in the same direction when using full stimulus onset to offset reaction times ( $ps < 0.01$ ). **C.** Interaction between cue and instruction on accuracy. As with reaction times, accuracy rates did not differ between push and pull trials after being cued to reverse instructions, whereas accuracy was significantly higher on pull trials following a “same” cue

**Table 1** Neural foci consistent with a significant activation to same-reverse cue (cue phase)

Region	BA	Volume (voxels)	x	y	z	F	<i>p</i>
Right precentral gyrus	4	95	37	-17	49	33.64	<0.00001
Right cingulate gyrus/SMA	6	53	12	-26	66	30.49	0.000015
Right superior temporal gyrus	22	50	52	-9	6	34.62	<0.00001
Left superior temporal gyrus	22	38	-56	-8	5	31.08	0.000013
Left lentiform nucleus		19	-23	5	-5	39.39	<0.00001
Left Medial PFC/ pre-SMA	6/32	18	-11	-21	68	27.32	0.00003

BA Brodmann Area; PFC prefrontal cortex; SMA supplementary motor area; x,y,z, F, p: MNI coordinates, F and *p* values for peak voxel

activation, we conducted six separate mixed-effect linear models applied to participants' reaction times with only one of each cluster's trial-level activation as predictor. For each cortical clusters (5 of 6 regions), cue activation was negatively related to subsequent reaction time to stimulus onset, with standardized linear coefficients ranging from  $-0.070$  to  $-0.094$  ( $ps < 0.001$ ; ensuring FDR Benjamini-Hochberg correction for multiple comparisons at a family-wise error rate/FWER of  $p < 0.05$ ; see Table S.3. in *Supplementary Material*). Thus, the more activation during the cue phase ("same"), the faster participants were in initiating movement in response to the upcoming stimulus.

### Stimulus phase

**Main effect of cue** Two areas in the right middle frontal gyrus/IFG (BA 47; 18 voxels; Peak Voxel MNI Coordinates: 41,42,-8;  $F = 25.1$ ,  $p < 0.00001$ ) and the right inferior parietal lobule (BA 40; 16 voxels; Peak Voxel MNI Coordinates: 42,-56,47;  $F = 26.7$ ,  $p < 0.00001$ ) were identified, in which activation was higher following a "reverse" cue relative to "same" cue (Fig. 3A-B). Activation to stimulus onset a following "reverse" cue in these areas was negatively related to trait negative affect based on PANAS scores ( $r = -0.58$ ,  $p = 0.003$  for IFG and  $r = -0.50$ ,  $p = 0.012$  for IPL; surviving FDR Benjamini-Hochberg correction at a FWER of  $p < 0.05$ ; Fig. 3C), but not with positive affect PANAS scores ( $ps > 0.05$ ).

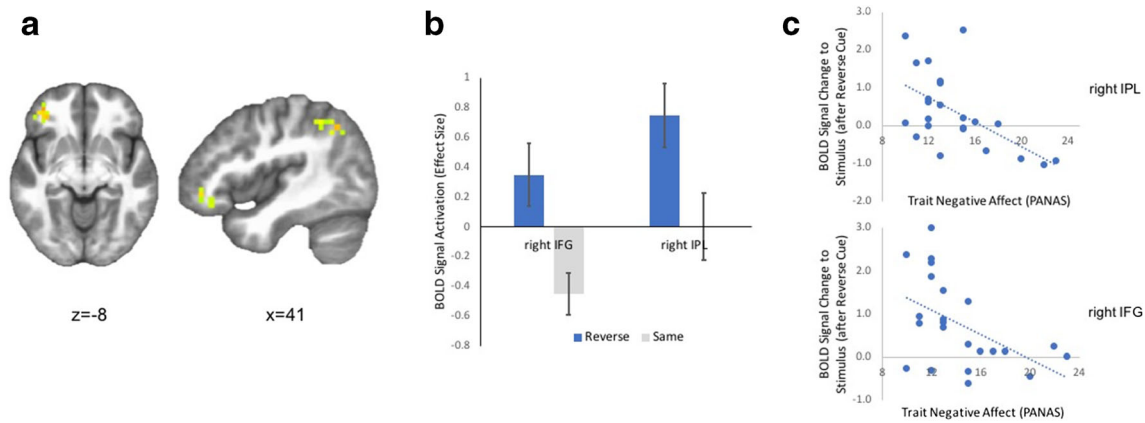
To gauge the nature of the relationship between pre-emptive cognitive control processes and subsequent stimulus reactive processes, we conducted an additional descriptive analysis. Specifically, we investigated the relationship between the extent of prestimulus deactivation in the six regions previously identified and stimulus-based activation in the IFG and IPL regions on reverse cue trials. Stimulus activation following a reverse cue in the right IPL was positively correlated with reverse cue activation (cue phase) in the left superior temporal gyrus ( $r = 0.59$ ,  $p = 0.003$ ), as well as in the right cingulate gyrus/SMA ( $r = 0.54$ ,  $p = 0.007$ ), the left medial PFC/pre-SMA ( $r = 0.53$ ;  $p = 0.008$ ), the right superior temporal gyrus ( $r = 0.51$ ,  $p = 0.013$ ), and the right precentral gyrus ( $r = 0.49$ ,  $p = 0.018$ ; surviving FDR Benjamini-

Hochberg correction at a FWER of  $p < 0.05$ ). Other ROI correlations did not reach statistical significance.

**Modulation of valence by cue** One cluster in the right precentral gyrus (BA 4; cluster volume = 33 voxels; Peak Voxel MNI Coordinates: 36,-22,37;  $F = 21$ ,  $p < 0.0001$ ; Fig. 4A) exhibited activation consistent with significant valence X cue interaction at stimulus presentation. In this region, higher activation to happy relative to angry faces was observed during trials following "same" cues, while such direction of activation was reversed and smaller on trials following "reverse" cues (Fig. 4B). Looking more closely at these neural patterns by instruction type, activation was significantly higher on push trials involving a happy face following "same" cues relative to all other conditions, which appeared to drive the two-way interaction effect reported above (Fig. 4C). This is consistent with a three-way interaction of cue X valence X instruction, which was statistically significant in this region at a lower voxel-wise threshold ( $p < 0.005$ ). In contrast to happy face (positive valence) trials, a significant deactivation, of smaller amplitude than push-positive trials activation, was observed on pull-negative (angry face) trials following "same cues." Such deactivation was not significantly different from zero on reverse angry face trials ( $ps > 0.05$ ). No statistically significant relationship between such activation coefficients and negative or positive affect was observed.

To further describe the relationship between this push-positive activation contrast in the right precentral gyrus and cue phase neural activity, we conducted supplementary correlational analyses. After correction for multiple comparisons, we observed positive correlations between cue type activation during cue phase (Same – Reverse) and Same-Reverse stimulus activation on trials requiring a push response to positive faces (Same Push Positive – Reverse Pull Positive). These ranged from 0.58 to 0.74 ( $ps < 0.004$ ; surviving FDR Benjamini-Hochberg correction at FWER of  $p < 0.05$ ).

**Modulation of instruction by cue** No cluster survived correction for multiple comparison for the instruction X cue interaction.



**Fig. 3** **A.** BOLD signal associated with a significant main effect of cue on stimulus response activation in the right inferior frontal gyrus (IFG) and right inferior parietal lobule (IPL). **B.** BOLD signal activation (linear coefficient  $t$  statistic) by cue type (blue = reverse; grey = same) for each region. Significantly larger positive activations were observed on trials following “reverse” cues, relative to trials following “same” cues ( $p <$

0.001; error bars = SEM). **C.** Scatter plots representing the negative correlations observed between trait negative affect (measured with PANAS) and neural activation to stimulus following reverse cues in right IPL (top) and right IFG (bottom). Negative affect was negatively correlated with IPL ( $r = -.50$ ,  $p < 0.05$ ) and IFG ( $r = -.58$ ,  $p < 0.01$ ) recruitment

**Interaction between cue, instruction, and valence** While no clusters survived multiple comparison correction with a voxel-wise threshold of  $p < 0.001$ , one cluster in the right precentral gyrus survived a correction at voxel-wise threshold of  $p < 0.005$ , which overlapped with the cluster identified for the cue X valence interaction. As noted above, the pattern of activation was consistent with a three-way interaction.

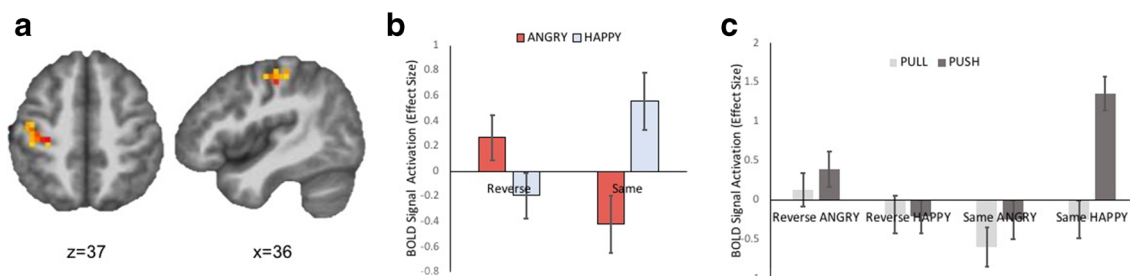
incongruent behavioral-social cues. To investigate this question, we developed a modified version of the standard Approach-Avoidance Task (AAT), in which participants were cued prior to each trial on whether they should reverse or not reverse the upcoming behavioral instruction to approach (i.e., pull joystick) or avoid (i.e., push joystick). We were particularly interested in the neural correlates of such rule reversal cueing before and during the reactive processing of the instructed action and valence of the social stimulus.

## Discussion

The goal of this study was to assess the role of proactive cognitive control on action tendencies of approach/avoidance in response to emotionally salient social stimuli (happy or angry faces). Specifically, we wanted to assess whether pre-emptive engagement of cognitive control may facilitate the processing of and behavioral adjustment to

### Prestimulus neural impact of cue

Cueing participants that they should “reverse” the subsequent instruction to either push or pull the joystick in response to a happy or angry face was associated with overall deactivation of a fronto-temporal network, including the precentral gyrus, the dorsal ACC, supplementary motor areas, and bilateral



**Fig. 4** **A.** BOLD signal associated with a significant interaction of cue and valence on stimulus response activation in the right precentral gyrus (Brodmann Area 4). **B.** BOLD signal activation (linear coefficient  $t$  statistic) in the right precentral gyrus by cue type and stimulus valence (red = negative/angry face; blue = positive/happy face). Significantly larger positive activations were observed on positive valence (happy face) relative to negative valence (angry face) trials following “same” cue (trials *not* requiring to reverse instruction;  $p < 0.01$ ). In contrast, activation on trials following “reverse” cues was higher for negative relative to positive

stimulus trial ( $p < 0.05$ ), but activations were smaller than “same” trials and not significantly different from zero ( $p > 0.05$ ; error bars = SEM). **C.** BOLD signal activation (linear coefficient  $t$  statistic) in the right precentral gyrus by cue type, stimulus valence, and action instruction (push vs. pull). This bar graph shows that the cue X valence interaction was driven by a three-way interaction, with significantly higher positive activation on push-happy trials following “same” cues relative to all other conditions ( $p < 0.05$ ; error bars = SEM)



superior temporal gyrus. Deactivation in the left lentiform gyrus also was observed in response to “reverse” relative to “same” cue. Interestingly, cued deactivation of these fronto-temporal regions generally appears to predict longer (less impulsive) reaction times upon stimulus presentation. Thus, such decrease in activation is associated with more careful behavioral adjustment in response to the social stimulus, presumably to minimize error (i.e., executing the opposite action). Specifically, premotor areas (BA 6) and anterior cingulate cortex have been shown to activate parametrically to the neural computations of goal-directed actions (Fincham, Carter, Van Veen, Stenger, & Anderson, 2002; Gardini et al., 2016), while superior temporal regions have been robustly involved in visuo-spatial processing supporting motor planning (Hanakawa, Dimyan, & Hallett, 2008; Shah-Basak, Chen, Caulfield, Medina, & Hamilton, 2018). Overall, this is consistent with the down-regulation of a spatial processing and motor planning neural network likely to track the mental set of the prepotent stimulus-response behavior. This preemptive deactivation may in turn lead to more careful, goal-directed performance.

These findings are congruent with the inhibitory control literature suggesting deactivations proportional to prediction errors in fronto-temporal areas during inhibitory control paradigms (Harlé et al., 2014; Harlé, Zhang, Ma, Yu, & Paulus, 2016) may provide a flexible neural inhibition system to help fine-tune behavior and decrease the chance of making a commission error. That is, while the present paradigm does not require actual prediction of the instruction type (i.e., the action to be performed), the prestimulus cue, rather than learned dynamic expectations, may provide a similar albeit more explicit guidance to facilitate subsequent performance. Relatedly, the deactivation of this network in the face of a reverse cue could reflect a more goal-oriented task-switching process, with pre-central motor areas playing an important role in both cognitive control and task-switching (Obeso, Robles, Muñoz-Marrón, & Redolar-Ripoll, 2013). Nonetheless, the observed deactivations of these areas in response to a reverse cue are consistent with disengagement or inhibition of a prepotent stimulus-response set.

### Overall effect of cue on stimulus-based neural recruitment

Stronger neural recruitment in the right IFG and right IPL was observed during the stimulus phase following a “reverse” cue relative to “same” cue. This is consistent with reactive engagement of inhibitory and attentional control neural areas following a reverse cue, i.e., when one has been prepared to reverse motor instructions. Indeed, the right IFG has been robustly identified as a key cognitive control region in various paradigms involving inhibition of learned, prepotent motor responses, including the stop-signal and go-no-go tasks (Aron, Behrens, Smith, Frank, & Poldrack, 2007; Aron, Fletcher,

Bullmore, Sahakian, & Robbins, 2003; Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010). The IPL is involved in implementing attentional (Shomstein, 2012) and cognitive (Garavan, Ross, Murphy, Roche, & Stein, 2002; Garavan, Ross, & Stein, 1999) control, and the right IPL in particular when conflict is present (Milham et al., 2001), e.g., as in incongruent stimulus-action instructions. Recent work suggests that these regions are specifically implicated in the computations of predicting action-response contingencies (Harlé et al., 2016; Liljeholm, Wang, Zhang, & O'Doherty, 2013), highlighting their role in motor inhibitory learning within uncertain environments.

Interestingly, on reverse instruction trials, smaller deactivations of these fronto-temporal regions, particularly the superior temporal gyrus, during the cue phase was associated with larger activation of the IPL at stimulus onset. One interpretation is that less efficient preemptive inhibition of these motor planning areas when cued to reverse instruction is associated with greater need to reactively recruit cognitive control neural regions in response to the stimulus. That is, IPL and IFG recruitment in response to stimulus (vs. cue phase) is consistent with the role of these regions in reactive cognitive control (Aron, 2011; Braver, 2012). Here, cueing may help to modulate such reactive control in a more indirect way, i.e., by putting the brakes on the neural network supporting prepotent response, in contrast to ramping up activation of a proactive cognitive control network.

It is noteworthy that trait negative affect may be associated with a reduced reactive recruitment of cognitive control regions following a reverse cue. Thus, while negative affect was not significantly related to cue-related preemptive inhibition of motor planning areas activation, we did observe a negative relationship between negative affect and reactive activation of cognitive control areas at stimulus onset. In other words, negative affect, which can be conceptualized as a type of cognitive load (Brinker, Campisi, Gibbs, & Izzard, 2013), may downregulate recruitment of this fronto-parietal network when it is particularly needed to prevent erroneous motor responses. Consistent with our findings, a recent study found that depressed mood was associated with impairment in reactive control involving emotional face cues, while proactive control and reactive control to neutral cues was relatively unaffected (Saunders & Jentsch, 2014). At the neural level, however, depression severity was found to attenuate event-related potentials (ERP) amplitude associated with both proactive and reactive control (West et al., 2010). Increased reactive control and impaired proactive control coupled with reduction in prefrontal recruitment has also been observed in anxiety (Fales, Barch, Burgess, et al., 2008a). This contrasting finding to depressive/negative mood studies is consistent with the moderating effect of heightened arousal in anxiety and its deleterious effect on prefrontal function (Arnsten, 2009), independently of biases related to negatively valenced

cognitions (as in sad mood or depression). This raises the possibility of important individual differences and distinct valence versus arousal dimensional effects in proactive control recruitment vs reactive processing within negative affective pathology. For instance, anxious arousal with relatively mild negative affect may make one more susceptible to proactive control impairments and cue-related neural differences at the prestimulus stage in the present paradigm, while depressive/negatively valenced mood may be primarily associated with reactive control alterations, as observed here. While our sample was relatively healthy, the observed negative affect range captured by the PANAS may have been more reflective of depressive negative affect than anxiety or anxious arousal. We also did not observe any relationship between positive affect and cognitive control in the present study. While others have investigated the effect of positive affect and reward sensitivity on cognitive control (Chiew & Braver, 2014; Dreisbach & Goschke, 2004; Fröber & Dreisbach, 2012; Van Steenbergen, Band, & Hommel, 2009; van Steenbergen, Eikemo, & Leknes, 2019), these studies have primarily relied on an experimental manipulation of affect. We surmise that the absence of positive affect results in the present study may be due to our relatively small and affectively homogenous sample, which resulted in a relatively small range of positive affect characteristics. Moreover, unlike this other work, the present study did not manipulate mood (e.g., with psychological or pharmacological induction), but rather assessed the relationship between cognitive control and individual differences in trait affect based on self-report. Such experimental affective manipulation may be more robust to uncover these interactive effects. More research is needed to disambiguate the role of affect, including different types of negative and positive emotionality, on the cognitive control processes regulating approach-avoidance tendencies.

#### **Effect of proactive cognitive control on neural response to stimulus valence-action incongruence**

Cue was found to moderate the relationship between action instruction and valence neural activations, in that being instructed to push (away) a happy face following a “same” cue (i.e., not necessitating to reverse the instruction) was associated with significantly higher activation in the right precentral gyrus. In contrast, this region was not significantly activated in other valence-instruction combinations following a “same” cue, or in any condition following a “reverse” cue. This motor area overlaps with a cluster identified as significantly deactivated in response to reverse cues during the cue phase (see first region in Table 1). This specific region of the primary motor cortex is involved in encoding the direction, speed, and extent of planned movement and has extensive structural and functional connections to other fronto-parietal areas, including pre-SMA (BA 6) and somatosensory cortex

(Guye et al., 2003; Mars et al., 2009). It also receives input from posterior parietal cortex, including superior and inferior parietal lobule in relation to attentional control and movement planning (Desmurget et al., 1999; Guye et al., 2003). Thus, engagement of this area is consistent with greater recruitment of neural resources associated with the execution of task-based actions, which activates to implement an incongruent avoidance action in response to a positive cue, but is down-regulated when cued to reverse instructions.

This finding is first in line with research showing more recruitment of cognitive control and longer reaction times among healthy individuals when having to push/avoid a positively valenced stimulus in the AAT (Bertsch et al., 2018; Ernst et al., 2013) and in other approach-avoidance paradigms (Saraiva, Schüür, & Bestmann, 2013). Thus, during “same” cue trials, i.e., in the absence of proactive inhibition of motor planning areas, avoiding a positive face appears to prompt greater motor cortical recruitment to execute an action. Although we should note that, behaviorally, no such interaction reached statistical significance, e.g., in terms of reaction times. In contrast, a “reverse” cue before the stimulus onset was associated with overall lower activation, which did not distinguish action-valence conditions. That is, having to push a happy stimulus after a reverse cue (i.e., on reverse-pull trials) does not result in the same increased precentral gyrus activation. Interestingly, greater differential activation to same versus reverse cue (during the cue phase) was subsequently associated with greater differential activation to pushing a happy face under same versus reverse condition. In other words, greater cued inhibition of a motor planning and execution regions was associated with greater reduction in the recruitment of the right precentral gyrus to incongruent happy trials in a reverse instruction context. While these are only correlational analyses, one possible explanation is that being preemptively cued to reverse action instructions may eliminate such automatic behavioral tendency towards a positive social cue. That is, proactive inhibition of neural areas supporting the prepotent motor response set may reduce the need of a more reactive motor recruitment to execute an incongruent avoid-positive response. Alternatively, in light of slower reaction times to approach/avoid stimuli following reverse instruction cues, this finding may reflect a more general increase in cognitive resources being deployed following a reverse cue, which may in turn impair attention to and neural reactivity to contextual congruency. Thus, being cued to reverse instruction in the task may have a more indirect effect impact on the reduced engagement of reactive control to incongruent contexts (e.g., avoiding a happy face).

The present study has several limitations, including a small sample size, and the fact that our sample was relatively healthy with limited range of affective characteristics. Moreover, in the absence of reversal cue, a congruency effect on reaction times was only observed for positive faces, which limits the

interpretation of potential congruency bias for angry faces in the present study. However, we note that such behavioral effects on motor response times are not reliably found particularly for angry faces, which can be more ambiguous to process and motivate (Phaf et al., 2014), and also may be sensitive to participant's supine position in the fMRI scanner environment. This does not preclude, however, reliable approach-avoidance congruency effects at the neural processing level (Derntl et al., 2011; Ernst et al., 2013; Korucuoglu et al., 2014; Roelofs et al., 2009). We also note that, as with any interference paradigm with uneven cue type ratios, an uncertainty or surprise effect cannot be fully ruled out and may partially contribute to the observed effects, particularly during the cue presentation phase. However, the present task design and analyses allowed to differentiate cue presentation from the face/border stimuli presentation. Results pertaining to reactive control and their relationship to affective measures, bound to this distinct stimulus phase, are thus less likely to be confounded by cue frequency effects (e.g., surprise, differential effort). Finally, in the present paradigm, the cue signaled a reversal of instruction rather than preparing for a specific action (as seen more commonly in proactive-reactive control paradigms). This approach, however, allowed us to focus our analyses on a broader form of pre-stimulus proactive control, independent of the behavioral activation direction of approach vs avoidance. This approach may also have more ecological validity for developing cognitive bias modification training to boost socio-affective cognitive flexibility in individuals with negative affect.

One final point is that the present study emphasis was on the social-affective nature of approach-avoidance bias, using human faces as target stimuli. While socially relevant stimuli, such as emotional faces, are likely to evoke more reliable approach-avoidance biases across individuals, future research is needed to determine whether the present findings extend to more visually and psychologically complex stimuli (e.g., emotionally evocative scenes) as they may be more ambiguous and may require greater information processing. Finally, while faces may have some universal effect on approach-avoidance biases across humans, the gender and/or ethnicity of the faces being processed may moderate individuals' responses based on their own gender and ethnicity. Investigating such effects and how it further modulates the recruitment of cognitive control in paradigms, such as the CAAT will be important for future research.

## Summary and conclusions

Overall our findings are consistent with the notion that proactively engaging cognitive control (e.g., by cuing to reverse task instruction) helps deactivate a prepotent motor set execution system, which may in turn reduce the need to recruit a more reactive form of cognitive control, particularly when

faced with incongruent stimulus-action pairings. While these results will have to be replicated in larger samples, they suggest that reactive recruitment of cognitive control brain regions in a cued inhibition context may be particularly challenging for those individuals with higher trait negative affect. In contrast, negative affective pathology did not significantly modulate cued inhibition of brain regions associated with visual-spatial and motor processes in the present cued AAT paradigm, which may reflect a more proactive form of cognitive control. This type of cognitive control may be helpful in adapting to and downregulating incongruent valence-action tendencies in a social context, including among individuals with negative affective pathology.

**Acknowledgements** This research was supported by the Veterans Health Administration IK2CX001584, CX001600, and IK2CX000864 awarded to Katia Harlé, Jessica Bomyea, and Andrea Spadoni, respectively, as well as Merit I01-CX001542 and I01-CX000715, awarded to Alan Simmons, and NIMH funding (R00MH090243) awarded to Charles Taylor. Neither of the experiments reported in this article was formally preregistered. Neither the data nor the materials have been made available on a permanent third-party archive; requests for the data or materials can be sent via email to the lead author. None of the authors have any conflicts of interest to declare.

## References

- Ansari, T. L., & Derakshan, N. (2011). The neural correlates of impaired inhibitory control in anxiety. *Neuropsychologia*, *49*, 1146–1153.
- Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, *10*, 410–422.
- Aron, A. R. (2011). From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biological Psychiatry*, *69*, e55–e68.
- Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *The Journal of Neuroscience*, *27*, 3743–3752.
- Aron, A. R., Fletcher, P. C., Bullmore, T., Sahakian, B. J., & Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, *6*, 115–116. <https://doi.org/10.1038/nm1003>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
- Becker, D., Jostmann, N. B., Wiers, R. W., & Holland, R. W. (2015). Approach avoidance training in the eating domain: Testing the effectiveness across three single session studies. *Appetite*, *85*, 58–65.
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*, *37*, 90–101.
- Bertsch, K., Roelofs, K., Roch, P. J., Ma, B., Hensel, S., Herpertz, S. C., & Volman, I. (2018). Neural correlates of emotional action control in anger-prone women with borderline personality disorder. *Journal of Psychiatry & Neuroscience: JPN*, *43*, 161.

- Bomyea, J., Taylor, C. T., Spadoni, A. D., & Simmons, A. N. (2018). Neural mechanisms of interference control in working memory capacity. *Human Brain Mapping, 39*, 772–782.
- Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences, 16*, 106–113.
- Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences, 106*, 7351–7356.
- Brevers, D., Bechara, A., Kilts, C., Antoniali, V., Bruylant, A., Verbanck, P., ... Noël, X. (2018). Competing motivations: Proactive response inhibition toward addiction-related stimuli in quitting-motivated individuals. *Journal of Gambling Studies, 34*, 785–806.
- Brinker, J. K., Campisi, M., Gibbs, L., & Izzard, R. (2013). Rumination, mood and cognitive performance. *Psychology, 4*, 224.
- Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology, 65*, 5.
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin, 25*, 215–224.
- Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective, & Behavioral Neuroscience, 14*, 509–529.
- Chiu, Y.-C., Cools, R., & Aron, A. R. (2014). Opposing effects of appetitive and aversive cues on go/no-go behavior and motor excitability. *Journal of Cognitive Neuroscience, 26*, 1851–1860.
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*, 162–173. <https://doi.org/10.1006/cbmr.1996.0014>
- Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *Journal of Neuroscience, 29*, 11993–11999.
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System®(D-KEFS®): Examiner's Manual: Flexibility of Thinking, Concept Formation, Problem Solving, Planning, Creativity, Impulse Control, Inhibition*: Pearson.
- Derntl, B., Seidel, E.-M., Eickhoff, S. B., Kellermann, T., Gur, R. C., Schneider, F., & Habel, U. (2011). Neural correlates of social approach and withdrawal in patients with major depression. *Social Neuroscience, 6*, 482–501.
- Desmurget, M., Epstein, C., Turner, R., Prablanc, C., Alexander, G., & Grafton, S. (1999). Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nature Neuroscience, 2*, 563.
- Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 343–353. <https://doi.org/10.1037/0278-7393.30.2.343>
- Eisma, M. C., Rinck, M., Stroebe, M. S., Schut, H. A., Boelen, P. A., Stroebe, W., & van den Bout, J. (2015). Rumination and implicit avoidance following bereavement: an approach avoidance task investigation. *Journal of Behavior Therapy and Experimental Psychiatry, 47*, 84–91.
- Ernst, L. H., Ehlis, A.-C., Dresler, T., Tupak, S. V., Weidner, A., & Fallgatter, A. J. (2013). N1 and N2 ERPs reflect the regulation of automatic approach tendencies to positive stimuli. *Neuroscience Research, 75*, 239–249.
- Fales, C., Barch, D., Burgess, G., Schaefer, A., Mennin, D., Gray, J., & Braver, T. (2008a). Anxiety and cognitive efficiency: differential modulation of transient and sustained neural activity during a working memory task. *Cognitive, Affective, & Behavioral Neuroscience, 8*, 239–253.
- Fales, C. L., Barch, D. M., Rundle, M. M., Mintun, M. A., Snyder, A. Z., Cohen, J. D., ... Sheline, Y. I. (2008b). Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression. *Biological Psychiatry, 63*, 377–384.
- Fincham, J. M., Carter, C. S., Van Veen, V., Stenger, V. A., & Anderson, J. R. (2002). Neural mechanisms of planning: a computational analysis using event-related fMRI. *Proceedings of the National Academy of Sciences, 99*, 3346–3351.
- Fröber, K., & Dreisbach, G. (2012). How positive affect modulates proactive control: Reduced usage of informative cues under positive affect with low arousal. *Frontiers in Psychology, 3*, 265.
- Garavan, H., Ross, T., Murphy, K., Roche, R., & Stein, E. (2002). Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *Neuroimage, 17*, 1820–1829.
- Garavan, H., Ross, T., & Stein, E. (1999). Right hemispheric dominance of inhibitory control: an event-related functional MRI study. *Proceedings of the National Academy of Sciences, 96*, 8301–8306.
- Gardini, S., Venneri, A., McGeown, W. J., Toraci, C., Nocetti, L., Porro, C. A., & Caffarra, P. (2016). Brain activation patterns characterizing different phases of motor action: execution, choice and ideation. *Brain Topography, 29*, 679–692.
- Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *The Journal of Neuroscience, 31*, 7867–7875.
- Guitart-Masip, M., Huys, Q. J., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage, 62*, 154–166.
- Guye, M., Parker, G. J., Symms, M., Boulby, P., Wheeler-Kingshott, C. A., Salek-Haddadi, A., ... Duncan, J. S. (2003). Combined functional MRI and tractography to demonstrate the connectivity of the human primary motor cortex in vivo. *Neuroimage, 19*, 1349–1360.
- Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *Neuroimage, 50*, 1313–1319.
- Hanakawa, T., Dimyan, M. A., & Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cerebral Cortex, 18*, 2775–2788.
- Harlé, K. M., Shenoy, P., Stewart, J. L., Tapert, S. F., Yu, A. J., & Paulus, M. P. (2014). Altered Neural Processing of the Need to Stop in Young Adults at Risk for Stimulant Dependence. *The Journal of Neuroscience, 34*, 4567–4580.
- Harlé, K. M., Zhang, S., Ma, N., Yu, A. J., & Paulus, M. P. (2016). Reduced Neural Recruitment for Bayesian Adjustment of Inhibitory Control in Methamphetamine Dependence. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*, 448–459. <https://doi.org/10.1016/j.bpsc.2016.06.008>
- Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: the approach-avoidance task. *Behaviour Research and Therapy, 45*, 2990–3001.
- Holmes, A. J., & Pizzagalli, D. A. (2007). Task feedback effects on conflict monitoring and executive control: Relationship to subclinical measures of depression. *Emotion, 7*, 68.
- Jacobus, J., Taylor, C. T., Gray, K. M., Meredith, L. R., Porter, A. M., Li, I., ... Squeglia, L. M. (2018). A multi-site proof-of-concept investigation of computerized approach-avoidance training in adolescent cannabis users. *Drug and Alcohol Dependence, 187*, 195–204.
- Korucuoglu, O., Gladwin, T. E., & Wiers, R. W. (2014). Preparing to approach or avoid alcohol: EEG correlates, and acute alcohol effects. *Neuroscience Letters, 559*, 199–204.
- Krug, M. K., & Carter, C. S. (2010). Adding fear to conflict: a general purpose cognitive control network is modulated by trait anxiety. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 357–371.

- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377–1388.
- Lesh, T. A., Westphal, A. J., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., ... Carter, C. S. (2013). Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia. *NeuroImage: Clinical*, *2*, 590–599.
- Liljeholm, M., Wang, S., Zhang, J., & O'Doherty, J. P. (2013). Neural correlates of the divergence of instrumental probability distributions. *The Journal of Neuroscience*, *33*, 12519–12527.
- Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171.
- Loijen, A., Vrijens, J. N., Egger, J. I., Becker, E. S., & Rinck, M. (2020). Biased approach-avoidance tendencies in psychopathology: A systematic review of their assessment and modification. *Clinical Psychology Review*, 101825.
- Mars, R. B., Klein, M. C., Neubert, F.-X., Olivier, E., Buch, E. R., Boorman, E. D., & Rushworth, M. F. (2009). Short-latency influence of medial frontal cortex on primary motor cortex during action selection under conflict. *Journal of Neuroscience*, *29*, 6926–6931.
- Maxwell, J. S., & Davidson, R. J. (2007). Emotion as Motion Asymmetries in Approach and Avoidant Actions. *Psychological Science*, *18*, 1113–1119.
- Milham, M., Banich, M., Webb, A., Barad, V., Cohen, N., Wszalek, T., & Kramer, A. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Cognitive Brain Research*, *12*, 467–473.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.
- Najmi, S., Kuckertz, J. M., & Amir, N. (2010). Automatic avoidance tendencies in individuals with contamination-related obsessive-compulsive symptoms. *Behaviour Research and Therapy*, *48*, 1058–1062.
- Ng, J., Chan, H. Y., & Schlaghecken, F. (2012). Dissociating effects of subclinical anxiety and depression on cognitive control. *Advances in Cognitive Psychology*, *8*, 38.
- Obeso, I., Robles, N., Muñoz-Marrón, E., & Redolar-Ripoll, D. (2013). Dissociating the role of the pre-SMA in response inhibition and switching: a combined online and offline TMS approach. *Frontiers in Human Neuroscience*, *7*, 150.
- Phaf, R. H., Mohr, S. E., Rotteveel, M., & Wicherts, J. M. (2014). Approach, avoidance, and affect: a meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Frontiers in Psychology*, *5*, 378.
- Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*, 105–120.
- Rinck, M., Wiers, R. W., Becker, E. S., & Lindenmeyer, J. (2018). Relapse prevention in abstinent alcoholics by cognitive bias modification: Clinical effects of combining approach bias modification and attention bias modification. *Journal of Consulting and Clinical Psychology*, *86*, 1005.
- Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. *Boenningstedt: SmartPLS GmbH*.
- Roelofs, K., Elzinga, B. M., & Rotteveel, M. (2005). The effects of stress-induced cortisol responses on approach-avoidance behavior. *Psychoneuroendocrinology*, *30*, 665–677.
- Roelofs, K., Minelli, A., Mars, R. B., van Peer, J., & Toni, I. (2009). On the neural control of social emotional behavior. *Social Cognitive and Affective Neuroscience*, *4*, 50–58.
- Saraiva, A. C., Schüür, F., & Bestmann, S. (2013). Emotional valence and contextual affordances flexibly shape approach-avoidance movements. *Frontiers in Psychology*, *4*, 933.
- Saunders, B., & Jentsch, I. (2014). Reactive and proactive control adjustments under increased depressive symptoms: insights from the classic and emotional-face Stroop task. *The Quarterly Journal of Experimental Psychology*, *67*, 884–898.
- Seidel, E.-M., Habel, U., Finkelmeyer, A., Schneider, F., Gur, R. C., & Derrtl, B. (2010). Implicit and explicit behavioral tendencies in male and female depression. *Psychiatry Research*, *177*, 124–130.
- Shah-Basak, P. P., Chen, P., Caulfield, K., Medina, J., & Hamilton, R. H. (2018). The role of the right superior temporal gyrus in stimulus-centered spatial processing. *Neuropsychologia*, *113*, 6–13.
- Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: top-down and bottom-up attentional control. *Frontiers in Integrative Neuroscience*, *6*, 38.
- Swart, J. C., Frank, M. J., Määttä, J. I., Jensen, O., Cools, R., & den Ouden, H. E. (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biology*, *16*, e2005979.
- Taylor, C. T., & Amir, N. (2012). Modifying automatic approach action tendencies in individuals with elevated social anxiety symptoms. *Behaviour Research and Therapy*, *50*, 529–536.
- Van Steenbergen, H., Band, G. P., & Hommel, B. (2009). Reward Counteracts Conflict Adaptation Evidence for a Role of Affect in Executive Control. *Psychological Science*, *20*, 1473–1477.
- van Steenbergen, H., Eikemo, M., & Leknes, S. (2019). The role of the opioid system in decision making and cognitive control: A review. *Cognitive, Affective, & Behavioral Neuroscience*, *19*, 435–458.
- van Uijen, S. L., van den Hout, M., & Engelhard, I. M. (2015). Active approach does not add to the effects of in vivo exposure. *Journal of Experimental Psychopathology*, *6*, 112–125.
- Volman, I., Roelofs, K., Koch, S., Verhagen, L., & Toni, I. (2011a). Anterior prefrontal cortex inhibition impairs control over social emotional actions. *Current Biology*, *21*, 1766–1770.
- Volman, I., Toni, I., Verhagen, L., & Roelofs, K. (2011b). Endogenous testosterone modulates prefrontal-amygdala connectivity during social emotional behavior. *Cerebral Cortex*, *21*, 2282–2290.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063.
- West, R., Choi, P., & Travers, S. (2010). The influence of negative affect on the neural correlates of cognitive control. *International Journal of Psychophysiology*, *76*, 107–117.
- Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*, 2020.
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychological Science*, *22*, 490–497.
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., & Strack, F. (2010). Retraining automatic action - tendencies to approach alcohol in hazardous drinkers. *Addiction*, *105*, 279–287.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.