Social Cognitive and Affective Neuroscience, 2019, 709–718

doi: 10.1093/scan/nsz045
Advance Access Publication Date: 3 July 2019
Original article

Social evaluations under conflict: negative judgments of conflicting information are easier than positive judgments

Hannah U. Nohlen,[1,2] Frenk van Harreveld,[1] and William A. Cunningham[2]

[1]Department of Social Psychology, University of Amsterdam, The Netherlands and [2]Department of Psychology, University of Toronto, M5S 3G3 Toronto, Ontario, Canada

Correspondence should be addressed to Hannah Nohlen, Department of Psychology, University of Amsterdam, 1018 WT Amsterdam, The Netherlands. E-mail: h.u.nohlen@uva.nl

Abstract

In the current study, we used functional magnetic resonance imaging to investigate how the brain facilitates social judgments despite evaluatively conflicting information. Participants learned consistent (positive or negative) and ambivalent (positive and negative) person information and were then asked to provide binary judgments of these targets in situations that either resolved conflict by prioritizing a subset of information or not. Self-report, decision time and brain data confirm that integrating contextual information into our evaluations of objects or people allows for nuanced (social) evaluations. The same mixed trait information elicited or failed to elicit evaluative conflict dependent on the situation. Crucially, we provide data suggesting that negative judgments are easier and may be considered the 'default' action when experiencing evaluative conflict: weaker activation in dorsolateral prefrontal cortex during trials of evaluative conflict was related to a greater likelihood of unfavorable judgments, and greater activation was related to more favorable judgments. Since negative outcome consequences are arguably more detrimental and salient, this finding supports the idea that additional regulation and a more active selection process are necessary to override an initial negative response to evaluatively conflicting information.

Key words: conflict; ambivalence; impression formation; cognitive control; social cognition; person perception; decision making; social judgment; evaluation

Every day we evaluate and interact with others across a range of situations. Because human behavior is complex, it is not uncommon that information we gather about others is marked by ambivalence, for example, when we perceive a person as cold but competent. Often, we are forced to resolve such evaluative conflict toward a favorable (e.g. collaborate with this person) or unfavorable judgment (e.g. do not collaborate). In the current article, we extend existing literature by investigating how the brain facilitates these social decisions by weighing evaluative information in line with affordances of the situation. Importantly, we provide data suggesting that negative judgments can be considered the easier, 'default' response when experiencing evaluative conflict.

Social evaluations are influenced by aspects of the task or situation that provide goals in relation to which a person or object is evaluated. Thereby, situational affordances facilitate flexible, nuanced evaluations (Cunningham & Zelazo, 2007; Cunningham *et al.*, 2007). In a recent study, we suggested that situational affordances can resolve evaluative conflict by prioritizing specific information; that is, we may judge someone positively in a specific situation despite knowing that the person also has negative features (Nohlen *et al.*, 2016). For example,

Received: 14 September 2018; **Revised:** 11 June 2019; **Accepted:** 17 June 2019

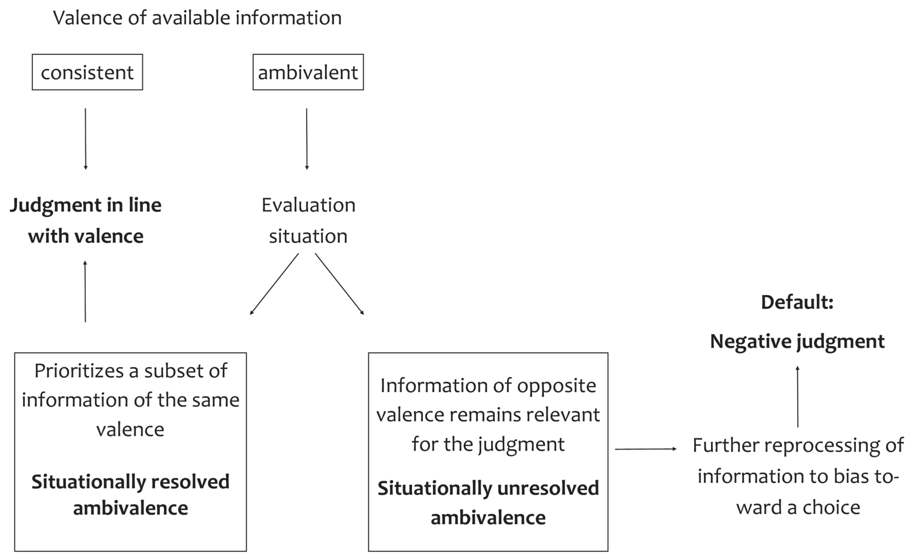

Fig. 1. Proposed processing of evaluative information in line with situational affordances.

we may judge a colleague who is charming and lazy positively when deciding whether to invite him or her to a social event because we prioritize the positive trait (charming) over the more negative one (lazy). Accordingly, affordances of the situation can also fail to resolve conflict if features of opposite valence remain relevant for the current judgment. For example, choosing whether to organize a social event with a friend who is charming and lazy makes both positively (i.e. charming) and negatively evaluated traits (i.e. lazy) important and evaluative conflict remains (Figure 1). These situations of evaluative conflict and specifically how judgments are facilitated despite evaluative conflict are the focus of the current study.

Cognitive conflict is traditionally studied in paradigms where a more salient, default response interferes with an objectively correct response (e.g. Stroop task and Eriksen Flanker task). Interestingly, there is no objectively correct response when conflict occurs between subjective evaluations, and the default response in these evaluative conflicts is unclear. We suggest that in such cases, negative judgments are the easier, default response based on two theoretical approaches. First, psychologists and neuroscientists have theorized that conflicts generate negative value (Botvinick, 2007; Braem et al., 2017; Dreisbach & Fischer, 2012; Inzlicht et al., 2015; Schouppe et al., 2015). For example, in a facial electromyographic study using a variant of the current impression formation paradigm, participants expressed less positive affect when ambivalent person information remained conflicted in the evaluation situation compared to when ambivalence was resolved (Nohlen et al., 2016). The negative value of conflict has also been suggested to extend to evaluations. Direct evidence comes from a study by Fritz & \Dreisbach (2013), who showed that presenting conflict (incongruent Stroop) primes increases the number of negative judgments of neutral targets (words or Chinese pictographs) that followed these primes.

Second, some have argued for a greater impact of negative information on evaluations because it arguably outweighs positive information in terms of salience and outcome consequences (Cacioppo *et al.*, 1997; Rozin & Royzman, 2001; Snyder & Tormala, 2017). For example, animals that are conflicted between approaching and avoiding a predator-infested water source tend to show avoidance behavior (Gray & McNaughton, 2000). This bias may be due to the fact that ignoring negative information can instill higher costs (e.g. being eaten) than ignoring positive information (e.g. drinking water; Ohman *et al.*, 2001). Similarly, negative judgments that insinuate avoidance in social situation (e.g. do not collaborate) could represent the 'safer choice' by maintaining the status quo when feeling torn between positive and negative evaluations (cf. Danziger *et al.*, 2011).

A tendency toward negative social judgments on the basis of conflicting person information can thus be expected because they may be easier due to the negative value of conflict which influences evaluations (e.g. Botvinick, 2007; Fritz & Dreisbach, 2013), and because negative information arguably outweighs positive information in terms of salience and outcome consequences (Gray & McNaughton, 2000; Rozin & Royzman, 2001).

If negative judgments are the easier response in evaluative conflicts, positive judgments should be more effortful and take more time. Supporting this argument, research has shown that even though ambiguous (i.e. surprised) facial expressions are primarily judged as negative, positive judgments become more likely when participants take more time to evaluate the stimulus (Kim *et al.*, 2003; Neta *et al.*, 2009; Neta & Tong, 2016). This suggests that overriding an initial negative response to interpret an ambiguous stimulus positively requires additional regulatory processes; we need to invest effort and actively attend to positive information (Kim *et al.*, 2004, 2003; Tottenham *et al.*, 2013).

Many models of control and executive functioning focus on the interaction of the anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (DLPFC) to resolve such complex decision situations (Alexander & Brown, 2015; Carter & van Veen, 2007; Sallet *et al.*, 2011). These domain-general regions are also more active in impression formation task when evaluations are updated with information that is inconsistent with prior information (Ma *et al.*, 2012; Ma *et al.*, 2011). Because of their engagement during many tasks, it has proven difficult to pinpoint the specific role of these regions. Interpretations of dorsal ACC's (dACC's) function thus vary from conflict monitoring (Botvinick, 2007) and detection (e.g. Carter & van Veen, 2007) to violation of expectancies (Somerville *et al.*, 2006), negative affect and pain (Shackman *et al.*, 2011) or comparisons of value outcomes, to name a few (Boorman *et al.*, 2013; Hare *et al.*, 2011; Kolling *et al.*, 2016; Rushworth *et al.*, 2011). Even though there is thus some debate on the specifics of dACC functioning (e.g. Ebitz & Hayden, 2016; Heilbronner & Hayden, 2016), it is widely agreed that dACC is engaged when tasks are more difficult and effortful as is, for example, the case when we have to form an integrative judgment from conflicting information (Nohlen et al., 2014). Recently, the Expected Value of Control theory (Shenhav et al., 2013) suggested that dACC has the role of a 'controller' that detects and signals an increased need for control through a cost–benefit analysis for optimal control allocation (Heilbronner & Hayden, 2016; Shenhav et al., 2013). Important in the current paradigm, these control-eliciting situations are often those in which one behavior, generally referred to as the default response, is suppressed in favor of another behavior that serves current goals better (Ebitz & Hayden, 2016; Kool et al., 2017).

According to this idea, difficult choice situations signaled by the ACC are relayed to the DLPFC, which is critically involved in the implementation of control (Botvinick et al., 2001). Specifically, DLPFC has been associated with biasing processing in line with salient goals, meaning that it is involved in attending to task-relevant information and selecting context-appropriate responses (Badre & Wagner, 2004; Miller & Cohen, 2001). Supporting this, Hughes et al. (2017) found both dACC and DLPFC to play a role in an intergroup social judgments task. A failure to engage these regions was associated with increased ingroup bias in that participants did not adjust their impression of ingroup members to incorporate negative information and stuck to their default (biased) response. Based on the idea that negative judgments are the default response in situations of evaluative conflict, positive judgments should represent a move away from the default response and thus be mediated by greater DLPFC and dACC activation.

Present research and hypotheses

The goal of our study was to investigate how the human brain facilitates social judgments when information is conflicting. Two aspects were central. First, we were interested in replicating our previous work showing the flexibility of ambivalent person evaluations with the idea that the same positive and negative person information elicits evaluative conflict in some but not in other situations (Nohlen et al., 2016). Second, we examined the role of dACC and DLPFC in biasing evaluations toward positive or negative judgments when evaluative conflict remains. As far as we know, this is the first study investigating whether blood oxygen level-dependent (BOLD) signal can be related to the valence of social evaluations.

We used a novel forced-choice task in which participants judged target persons described by consistent (positive or negative) or ambivalent (positive and negative) traits in different situations that either allowed for prioritizing some traits over others (conflict is resolved) or not (conflict remains unresolved; cf. Nohlen et al., 2016). If negative judgments of conflicting information are the default, we should find an interaction effect of the valence of social judgment (positive, negative) and the presence of evaluative conflict on BOLD signal. More specifically, dACC and DLPFC response should be stronger if participants experience evaluative conflict and override the default negative judgment and judge the person positively. Accordingly, weaker dACC and DLPFC activation should be related to negative judgments under evaluative conflict. If conflict does not have a negative value that influences evaluations, brain response to the valence of social judgments should be independent from conflict.

Methods

Participants

Participants were 20 adults (11 male, 9 female) in the age range of 18 to 29 years (M = 22.7; s.d., 2.60). Participants provided informed consent, had normal or corrected-to-normal vision, and had no history of neurological problems. Nineteen of 20 participants were right handed. All procedures were approved by the local ethical committee.

Design and procedure

Target persons. Two to 5 days before scanning, participants received descriptions of four male target persons that consisted of a list of traits and a short text to make the traits more memorable. One target was described by positive traits (friendly, charming, enthusiastic, intelligent), one by negative (dominant, jealous, lazy, dumb), and two by the combination of the two positive and two negative traits each (dominant, jealous, enthusiastic, intelligent; friendly, charming, lazy, dumb). The combinations of names and traits were counterbalanced across participants, and we used a pretest ($N = 34$; Supplementary Material S1) to ensure that the trait combinations were evaluated as positive, negative or ambivalent. Participants memorized the combinations of names and traits before coming to the laboratory and were verbally tested on recall during take-in. If participants were not able to recall the traits and names of the targets, they were given additional time to learn them. Because we did not want differences in knowledge of the name–trait combinations to introduce noise in the data, exposure to the pairs may have varied between participants. Post-relearning, all participants were able to quickly recall the combinations when prompted.

Evaluation situations. The four target persons were evaluated in 21 different situations (cf. Nohlen et al., 2016). All situations were combined with each of the four target persons (84 trials). Based on a pretest (S1), situations were selected that varied in the degree to which they resolved conflict between ambivalent trait information by prioritizing a subset of either positive or negative traits. In the pretest, we used a one-item adaptation of Priester and Petty's (1996) subjective ambivalence scale, which assesses experienced conflict by asking participants to evaluate the degree to which they experienced 'mixed feelings and/or thoughts' toward each target person (described by their specific traits) in each of the different evaluation situations on a scale ranging from 0 (not at all) to 100 (very much). Results were used to categorize the combinations of ambivalent target persons and evaluation situations as representing either situationally resolved ambivalent (2 target persons × 10 situations = 20 trials), or unresolved ambivalent judgments (2 target persons × 11 situations = 22 trials), resulting in three critical trial types based on the combination of person traits and evaluation situation (Figure 2B). Note that the situation did not have to resolve conflict when targets were described by positive or negative (i.e. consistent) traits; they elicited unconflicted judgments across all situations (2 target persons × 21 situations = 42 trials). Whether situations resolved conflict between ambivalent target information was dependent on the specific target; the same situation could thus be categorized as situationally resolved ambivalent for one target and unresolved for another (Supplementary Material S2 for all stimuli).[1] Additionally, a manipulation check was added to this study to test whether this categorization was successful (see Manipulation checks).

A

B

Trial valence	Person information	Examples of evaluation situations
Univalent: positive (consistent)	**X**: Friendly, charming, intelligent, enthusiastic	Would you trust X to do your tax declaration?
Univalent: negative (consistent)	**X**: Lazy, dumb, dominant, jealous	Would you want to study for an exam with X?
Ambivalent: Situationally resolved	***A***: friendly, charming, lazy, dumb; ***B***: intelligent, enthusiastic, dominant, jealous	Would you trust ***A*** to do your tax declaration?
Ambivalent: Situationally unresolved	***A***: friendly, charming, lazy, dumb; ***B***: intelligent, enthusiastic, dominant, jealous	Would you want to study for an exam with ***B***?

Fig. 2. (A) Timing sequence of an experimental trial. (B) Example trial of each trial type.

Social Judgment task

Functional magnetic resonance imaging social judgment task. In the scanner, participants indicated their judgment ('yes,' 'no') to each combination of target person and evaluation situation by pressing one of two buttons with their right index and middle finger. Choice labels were counterbalanced between participants. Participants completed one functional run of 84 randomly presented trials with a 5 s break after every 21 trials. Trials started with a fixation cross (500 ms), and stimuli were presented until participants responded. After the response, participants saw a white dot on a black screen with varying presentation times (min, 4000 ms; max, 6600 ms) before the next trial (Figure 2A). Responses were recoded so that yes responses represent positive judgments and no responses represent negative judgments of the person in that situation.

Manipulation checks.

Before scanning: assessing ambivalence toward target persons. To confirm that information about the target persons was indeed perceived as consistent or conflicted, we assessed experienced ambivalence toward the four target persons with the subjective ambivalence scale (Priester & Petty, 1996). This scale assesses psychological conflict with three items anchored with 'Toward this person I... have completely one-sided feelings/feel no conflict/feel no indecision' (0) and 'have mixed feelings/feel maximum conflict/maximum indecision' (100). Participants responded on a slider without numeric labels ($\alpha = 0.71$). Ambivalence was calculated by taking the mean of these questions.[2]

After scanning: assessing ambivalence toward each target person in each situation. To verify that we correctly categorized the combinations of person information and evaluation situation as conflicting or conflict resolved on the basis of pretest data, participants indicated their ambivalence toward each target person in each situation using the same scale (Priester & Petty, 1996; e.g. 'I have mixed feelings/feel conflict/feel indecision about collaborating with X'). Responses were given on a slider (0–100) without numeric labels ($\alpha = 0.93$).[3]

Magnetic resonance imaging data acquisition

Imaging was conducted with a 3.0-T Philips Achieva scanner at the Spinoza Centre for Neuroimaging in Amsterdam. Head motion was limited by placing foam inserts around the head inside the head coil. Stimuli were presented using E-Prime and projected onto a screen in the magnet bore, which participants could see through a mirror attached to the head coil. Functional data were obtained using T2*-weighted echo-planar imaging in one event-related run (84 trials). The first two dummy scans were removed to allow for equilibration of T1 saturation effects [time repetition (TR), 2 s; time echo (TE), 28 ms; voxel size, 3 × 3 × 3 mm; field of view (FOV), 240^2]. A high-resolution T1-weighted sagittal scan was collected as anatomical reference (TR, 9.56 s; TE, 4.6 s; voxel size, 1.2 × 1.2 × 1.2; FOV, 224^2).

Functional magnetic resonance imaging preprocessing

Data were preprocessed using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FMRIB's Software Library (FSL, www.fmrib.ox.ac.uk/fsl). Data were corrected for motion using FMRIB's Linear Image Registration Tool (MCFLIRT; (FSL's Motion Correction using FMRIB's Linear Image Registration Tool; Jenkinson *et al.*, 2002); the brain was extracted using FSL's Brain Extraction Tool (BET; Smith, 2002) and spatially smoothed using a Gaussian kernel (full width at half maximum [FWHM] = 5 mm). The four-dimensional data set was grand-mean intensity normalized by a single multiplicative factor. We applied a high-pass temporal filter (Gaussian-weighted least-squares straight line fitting, with $\sigma = 50$ s). Registration to standard space images was carried out using FLIRT (Jenkinson *et al.*, 2002; Jenkinson & Smith, 2001).

Results

Behavior analysis

Manipulation checks.

Before scanning: assessing ambivalence toward target persons.. Results on the subjective ambivalence scale confirmed that the combination of traits successfully created unconflicted (consistent) or conflicted evaluations of the target persons independent of evaluation situation. Participants experienced less conflict regarding persons described by positive or negative traits ($M_{consistent} = 11.03$, $SE_{consistent} = 2.24$) than toward those described by positive and negative traits [$M_{ambivalent} = 40.67$, $SE_{ambivalent} = 3.13$; $t(19) = -8.610$, $P < 0.001$, $r = 0.89$].

During scanning: response behavior. In line with the categorization, target persons described by positive traits were evaluated positively in 416 of 420 trials (99%) independent of evaluation situation, and negatively described targets were evaluated negatively in 399 of 420 trials (95%). Correspondingly, judgments were mixed when target persons were described by positive and negative traits; positive judgments were given in 433 of 840 trials (51.5%) and negative judgments in the remaining 407 trials (48.5%). Splitting ambivalent trials according to situational resolution of conflict showed that negative judgments were given in 181 of the 400 conflict-resolved ambivalent trials (45.3%) and in 226 of the 440 unresolved ambivalent trials (51.4%). This difference was significant in that participants judged ambivalent target persons negatively more often when conflict remained unresolved (51.4%) compared to when it was situationally resolved [45.3%; $F(1,19) = 8.09$, $P = 0.01$, $\eta_p{}^2 = 0.30$].

After scanning: assessing ambivalence toward each target person in each situation. We assessed experienced conflict toward target persons in each evaluation situation to verify the pretest-based categorization between situations that resolved conflict and those that did not. Results confirmed the expected main effect of trial type [consistent, situationally resolved ambivalent, situationally unresolved ambivalent; $F(1.27,24.14) = 33.13$, $P < 0.001$, $\eta_p{}^2 = 0.64$ (Greenhouse–Geisser corrected)]. Experienced ambivalence toward unconflicted trials ($M_{consistent} = 7.96$, $SE_{consistent} = 2.5$) was lower than toward conflict-resolved ambivalent trials ($M_{AmbResolv} = 23.16$, $SE_{AmbResolv} = 3.00$; $P < 0.001$), which was lower than ambivalence toward conflict-unresolved ambivalent trials ($M_{AmbUnresolv} = 31.15$, $SE_{AmbUnresolv} = 3.74$, $P < 0.001$). The pattern of results provides confidence in the categorization into three trial types by combining trait information with evaluation situations.

Decision times. The time it takes to make a judgment provides an indication of judgment difficulty. To confirm that judgments were more difficult when conflict was not situationally resolved and to verify that negative judgments are easier when experiencing conflict, we conducted a repeated-measures analysis of variance (ANOVA) comparing decision times between trial type (consistent, resolved ambivalent, unresolved ambivalent) and the judgment participants provided (positive, negative). Expectedly, trial type influenced decision times [$F(2,38) = 37.89$, $P < 0.001$, $\eta_p{}^2 = 0.66$]; participants were quickest to evaluate persons with consistent traits ($M_{consistent} = 2.73$ s, $SE_{consistent} = 0.13$ s) and somewhat slower when traits were conflicted but resolved by situational affordances ($M_{AmbResolv} = 3.84$ s, $SE_{AmbResolv} = 0.28$ s; $P < 0.001$). Judgments took longest when traits were conflicted and remained conflicted in the evaluation situation ($M_{AmbUnresolv} = 4.20$ s, $SE_{AmbUnresolv} = 0.31$ s; $P = 0.02$), suggesting that these were experienced as most difficult. As expected, there was no main effect of response behavior, and participants were equally quick in providing negative and positive judgments [$F(1,19) = 1.53$, $P = 0.23$].

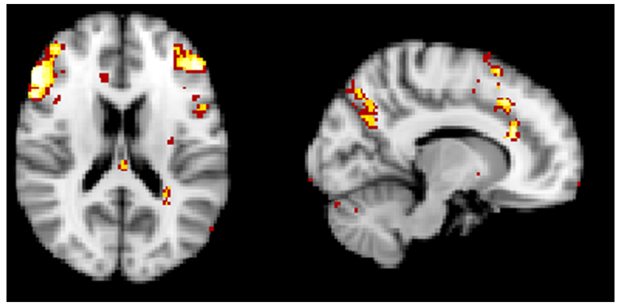

Fig. 3. Interaction effect of trial type and response behavior in the whole-brain MLM analysis (Montreal Neurological Institute [MNI] coordinates: $x = 12$, $y = -6$, $z = 18$). Activation thresholded at $P = 0.005$ (red) and $P = 0.001$ (yellow). The left side of the brain is shown on the right side and vice versa.

Critically, we found the expected interaction of response behavior and trial type [$F(2,38) = 6.54$, $P = 0.004$, $\eta_p{}^2 = 0.26$]. Confirming that negative judgments of conflicting stimuli are the easier response, decision times were quicker when participants judged situationally conflicting targets negatively ($M_{AmbUnresolvNeg} = 3.95$ s, $SE_{AmbUnresolvNeg} = 0.30$ s) than positively ($M_{AmbUnresolvPos} = 4.45$ s, $SE_{AmbUnresolvPos} = 0.34$ s; $P = 0.03$). In line with hypotheses, no difference in decision times between positive and negative judgments was found for ambivalent stimuli when situational affordances resolved conflict ($P = 0.24$). Regarding consistent stimuli, positive judgments were made faster ($M_{consistentPos} = 2.59$ s, $SE_{consistentPos} = 0.12$ s) than negative judgments ($M_{consistentNeg} = 2.88$ s, $SE_{consistentNeg} = 0.16$ s), $P = 0.03$.

Image analysis

Whole brain: multilevel model. Because we differentiated trials on the basis of subjects' judgments, we dealt with an unequal number of observations in each cell of the study design. To estimate the effect of trial type and subjects' judgment simultaneously, we thus constructed a multilevel model (MLM) that can deal with such unbalanced designs. Using AFNI's 3dDeconvolve function, we obtained β estimates for BOLD response magnitude for each voxel and trial by modeling functional magnetic resonance imaging (fMRI) time series with individual trial regressors based on onset times for stimulus presentation (https://afni.nimh.nih.gov/;-stim_times_IM; see Luttrell *et al.*, 2016; Mumford *et al.*, 2012; Stillman *et al.*, 2015 for similar approaches). Decision times were included as duration modulation. Trial type and response behavior were coded as categorical variables. Using R (R Core Team, 2013), we then modeled fMRI BOLD from the three trial types (consistent, resolved ambivalent, unresolved ambivalent), the judgments participants made (response behavior: positive *vs* negative judgment) and their interaction, with random intercepts on the subject level at each voxel to account for the repeated measures design (lme4-package; Bates *et al.*, 2015). Categorical variables were automatically dummy coded in the analysis, and in order to examine the effects of the two degrees of freedom simultaneously, we used the ANOVA function in the car library (car-package; Fox & Weisberg, 2019).

We observed a significant interaction between the effect of trial type and response behavior on brain activation in several regions, including bilateral DLPFC and dACC as hypothesized (Figure 3). For our analyses, we used an a priori threshold of $P < 0.005$ with a cluster size of 25 (Lieberman & Cunningham, 2009) to balance Type I and Type II errors. However, all effects

Table 1. Regions that showed a significant interaction effect of trial type and response behavior in the whole-brain MLM ($P < 0.001$, uncorrected)

Anatomical region	Laterality	Voxels	F-value	x	y	z
Dorsal PFC/IFG, BA 45	R	1406	38.7	18	78	46
	L	589	31.0	67	81	48
dACC/SFG, BA 6	L/R	816	32.7	44	72	62
Orbitofrontal cortex	R	215	27.6	25	75	33
Insula	L	104	32.2	59	75	35
Precuneous cortex	L/R	641	29.4	48	26	56
Posterior cingulate gyrus		40	20.4	44	48	49
Superior lateral occipital cortex	R	145	26.1	24	35	62
	R	98	30.5	28	28	65
	L	84	28.8	61	34	62
	R	32	23.2	29	24	58
Lateral occipital pole/visual cortex		53	20.7	59	13	31
Cerebellum	L	31	17.5	50	22	22
Premotor cortex, BA6	L	25	21.3	55	61	73

Notes. Only clusters that exceed a minimum size of 25 voxels are shown. Voxel coordinates of the maximally activated voxels are given. IFG, inferior frontal gyrus PFC, prefrontal cortex; SFG, superior frontal gyrus.

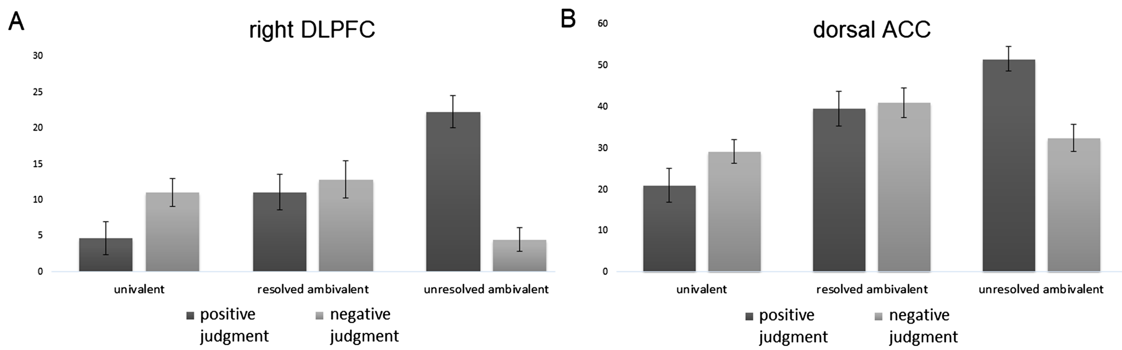

Fig. 4. Activation in right DLPFC (A) and dACC (B) by trial type and response behavior.

of interest were significant at a more stringent threshold and are more accessibly presented using a threshold of $P < 0.001$ with a cluster size of 25 (Table 1). For the ease of presentation, we use this more stringent threshold in the main paper for presentational purposes and place the full list of activations at the original threshold in the Supplementary Material S3.[4]

ROI analysis

Simple effects. *Post hoc* tests were conducted in SPSS on the DLPFC and dACC regions from the MLM analysis to decompose the interaction effect of trial type and response behavior. For each cluster, we extracted the β estimates for each trial and participant that were used to run the whole-brain MLM analysis. We averaged estimates across voxels for each trial and participant within each cluster. Note that the following analysis is applied to interpret the significant interaction effect by testing simple effects, not to run the same analysis on a subset of the data (Kriegeskorte *et al.*, 2010).

DLPFC. In line with expectations, we observed greater DLPFC activation for positive than negative judgments only when target persons were described by ambivalent traits and evaluative conflict remained unresolved by the situation in which the evaluation took place (right DLPFC, left DLPFC: $P < 0.001$; Figure 4A). When target traits were ambivalent but affordances resolved conflict, results showed no difference in DLPFC activation for positive and negative judgments (right DLPFC: $P = 0.67$; left DLPFC: $P = 0.82$). Activation in DLPFC was weaker when participants made positive compared to negative judgments when target traits were consistent (right DLPFC: $P = 0.009$; left DLPFC: $P = 0.02$).

dACC. The pattern of activation in the dACC mirrored DLPFC activation. When trials elicited (unresolvable) evaluative conflict, dACC activation was stronger when participants judged the person positively than negatively ($P < 0.001$; Figure 4B). We found no difference for judgment valence in dACC response to ambivalent target persons when affordances resolved conflict ($P = 0.79$). Finally, judging consistently described target persons negatively was associated with greater dACC activation than judging them positively ($P = 0.05$).

In a follow-up analysis, we verified that these effects cannot be ascribed to differences in decision times. Including decision times did not eliminate the critical interaction between trial type and response behavior on DLPFC activation [right: $\chi^2(2) = 22.63$, $P = 1.22e{-}05$; left: $\chi^2(2) = 21.95$, $P = 1.71e{-}05$; Type III Wald] and on dACC activation [$\chi^2(2) = 14.65$, $P = 0.0007$; see Supplementary Material S5 for details].

Explaining response behavior from variation in BOLD within trial type. A more stringent test of the hypothesis that dACC and DLPFC signal under evaluative conflict is related to going with or against the default evaluation is to test the relationship between judgment behavior and BOLD variability within different trial types while controlling for differences between participants and trial type. If we find that greater (weaker) DLPFC (dACC) activation is related to a greater likelihood of positive (negative) judgments—only within situationally unresolved ambivalent trials—we solidify the claim that positive judgments of conflicting

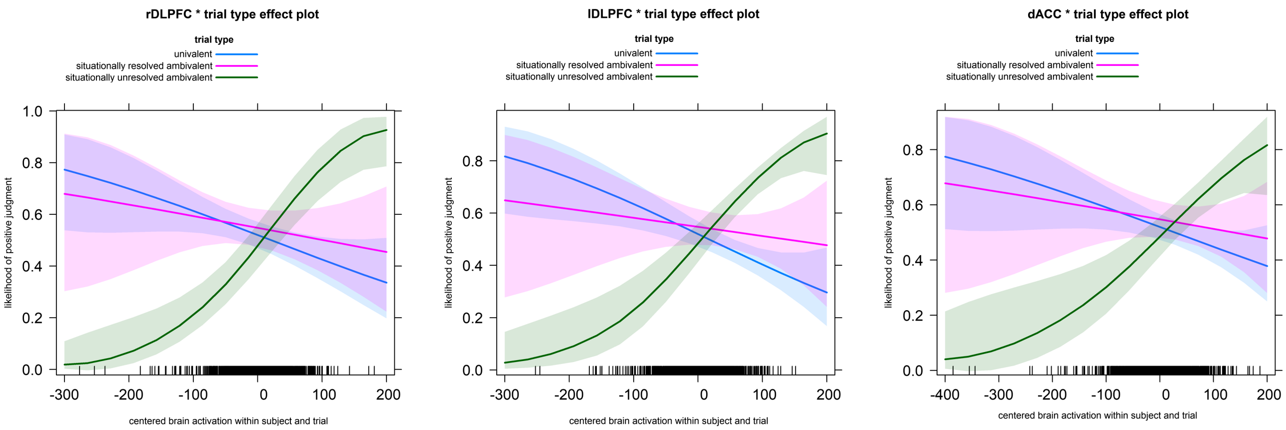

Fig. 5. Likelihood of positive and negative social judgments within each trial type on the basis of centered brain activation in right DLPFC (A), left DLPFC (B) and dACC (C). Shaded areas represent 95% confidence intervals.

information are more effortful. Compared to the previous analysis in which we found that positive judgments are in general related to greater DLPFC (dACC) response in situationally unresolved ambivalent trials across participants, this analysis looks for the relationship within participants and trial types. To test this, DLPFC (dACC) activation and decision times were centered within participant and trial type. Using R, trial type (consistent, resolved ambivalent, unresolved ambivalent), centered BOLD response and centered decision time effects, as well as their interaction with trial type, were modeled in a multilevel-model nested within subject. Trial type was coded as a categorical variable. Testing the interaction between centered BOLD response and trial type while controlling for decision times thus allows us to test whether response behavior can be explained by variation in BOLD signal in each type of trial. For readability, the discussion of the results focuses on the predicted interaction between BOLD response and trial type. Other effects are reported in the notes.

DLPFC[5]. Most critically, we found that variation in DLPFC activation predicted response behavior dependent on trial type [$\chi^2(2) = 22.16$, $P = 1.54e{-}05$; left DLPFC: $\chi^2(2) = 21.73$, $P = 1.91e{-}05$]. As can be seen in Figure 5, DLPFC activation had little effect on the valence of judgments when target persons' traits were consistent or the situation resolved ambivalence. When target traits were ambivalent and situational affordances did not resolve conflict, however, comparatively lower DLPFC activation was related to a greater likelihood of negative judgments and positive judgments were more likely under greater DLPFC activation. This supports our suggestion that negative judgments are easier when judging ambivalent targets, whereas greater DLPFC activation is necessary to bias toward more favorable judgments.

dACC[6]. The analysis was repeated for the dACC, again examining the interaction between centered dACC signal and trial type while controlling for decision times. The results mirrored the effect of DLPFC activation. Most importantly, variation in dACC activation within trial type condition predicted response behavior differently for different trial types [$\chi^2(2) = 14.92$, $P = 0.0006$; Figure 5C). However, the pattern was less clear than in DLPFC.

Discussion

In the current study, we investigated how the brain facilitates social judgments when person information is conflicting. We made two claims. First, we put forward that social evaluations are dynamic and have situational dependency (Cunningham *et al.*, 2007; Nohlen *et al.*, 2016). Second, we hypothesized that people have a tendency toward negative evaluations under evaluative conflict. To investigate this, our paradigm compared social judgments of people described by consistent or ambivalent trait information across different situations. Our findings confirm that social judgments of the same person vary across situations. Dependent on the situation in which the person was evaluated, ambivalent trait information elicited or failed to elicit evaluative conflict as represented in self-report, decision times and brain activation. Importantly, decision time and brain data converged in support of the hypothesis that negative evaluations are easier when experiencing evaluative conflict; participants were faster to make negative than positive judgments. Crucially, positive judgments were associated with greater DLPFC and dACC activation than negative judgments even when controlling for decision time, suggesting that differences in response magnitude are not simply an effect of processing time. Moreover, the variability of participants' DLPFC and dACC activation predicted the likelihood of positive and negative judgments only within evaluative conflict trials. The study thus provides converging evidence suggesting that saying no is easier than saying yes when experiencing evaluative conflict.

Our findings complement and extend prior work in a number of ways. Related work has focused on the negative value of response conflict, for example, showing that the negative value of conflict primes translates to negative judgments of following neutral targets (Fritz & Dreisbach, 2013). The current study extends this line of research showing that negative value of conflict translates to faster negative evaluations of the conflict stimulus itself. Additionally, whereas previous studies focused primarily on the negative value of response conflicts, we focused on evaluative conflict in the traditional sense, which requires the presence of conflict between positive and negative valence (i.e. ambivalence). Thereby, it provides a stringent test of the idea that negative evaluations of conflict-eliciting stimuli can be considered the default response, whereas positive evaluations require more control.

Important to note is that, even though we observed more negative judgments when participants experienced evaluative conflict compared to when they did not (on the basis of the same ambivalent person information), there was an equal number of positive and negative judgments within high conflict trials.

In line with others (Tottenham *et al.*, 2013), we suggest that overriding negative interpretations of ambiguous or conflicting information in favor of a positive interpretation requires regulatory processes. That is, even though negative judgments were easier, participants regularly seem to exert effort in favor of positive judgments. However, it is as yet unclear what motivates people to exert this effort and bias their judgment.

To our knowledge, this is also the first study that related the likelihood of positive or negative judgments of conflicting information to dACC and DLPFC activity. Negative judgments were made quicker, and weaker dACC and DLPFC response in trials eliciting evaluative conflict was linked to a greater likelihood of negative judgments, whereas stronger dACC and DLPFC response predicted a greater likelihood of positive judgments. Greater activation in dACC on trials eliciting evaluative conflict could be a prerequisite for implementing control and processing changes in DLPFC toward a more favorable judgment. However, the similar activation pattern observed here makes claims regarding their individual contribution challenging. Our interpretation is based on prior work suggesting that both regions play important roles in behavioral flexibility with dACC signaling the need for behavior adjustment, which is critically implemented by the DLPFC (e.g. Heilbronner & Hayden, 2016). In social evaluation tasks, regions in the lateral prefrontal cortex have been linked to updating initial impressions with incongruent information (Mende-Siedlecki *et al.*, 2013). For example, Bhanji and Beer (2013) demonstrated that modifying impressions toward a more favorable (i.e. positive) judgment is linked to parametrically increasing engagement of a region within lateral prefrontal cortex. Similar to our suggestion, they interpret this increased engagement as a reflection of cognitive effort. We argue that the engagement of DLPFC when changing one's social judgment is not linked to the valence (i.e. positivity) of the evaluation *per se* but dependent on whether the evaluation moves away from the default response. For example, Hughes *et al.* (2017) showed that engagement of DLPFC was related to incorporating negative information about positively evaluated ingroup members. Since it is not the valence of the evaluation but the movement away from the default response that engages DLPFC, and since positive evaluations of ingroup members are the default response, integrating negative information into the evaluation of an ingroup member requires more effort similar to moving toward a positive evaluation when experiencing evaluative conflict in our task. This may also suggest some interesting boundary conditions worthy of future investigation. Whereas we suggest that it is easier to evaluate individuals negatively based on conflicting information, there are likely situations under which this effect may be diminished or even reversed due to motives of the evaluator. Situational factors such as group membership, one's need to belong or stressful social situations (e.g. ostracism) may influence our interpretation and value judgment of certain traits as well as our global evaluation of another person. For example, when we feel socially rejected or lonely, we may be more inclined to evaluate another person positively despite conflicting trait information, thus shifting the default response to evaluatively conflicting information from negative to positive. Additionally, when evaluating ingroup and outgroup members on the basis of the same conflicting trait information, it may be relatively easier to judge an ingroup member positively because group membership carries positive value and we may process negative information about ingroup members in a biased manner (cf. Hughes *et al.*, 2017).

Usually, research into cognitive conflicts and the role of dACC and DLPFC in executive functioning is tested with non-social tasks in which perceptual or response conflict is correctly resolved by actively attending to and regulating behavior. Our study shows that the role of dACC and DLPFC in guiding behavior flexibly in more low-level cognitive control tasks also maps onto more complex tasks in which people have to search memory and weigh information in line with situational affordances. However, an important aspect of our study design was that participants were forced to provide a positive or negative judgment. By forcing a single-factor solution from evaluatively contradicting information, we created an evaluation situation in which neither response option fits participants' evaluation of the stimulus. Even though this approach mirrors many real-life situations, it prevents us from pulling apart possibly separate fMRI responses to evaluative conflicts and response conflicts. The distinction is less relevant for the current study since we were interested in the interaction between the presence of conflict and the judgment provided. However, independent of judgment valence, it may be that dACC response was partly driven either by the difficulty to choose between two unfitting response options or by the evaluative conflict that the stimulus elicits independently of choice. Future studies may investigate this further by comparing forced-choice situations with situations in which participants can indicate mixed evaluation toward an evaluatively conflicted stimulus.

Conclusions

The current study shows that social evaluations are dynamic and form in a particular situation, with a particular goal in mind. Complex and conflicted stimulus representations can be resolved by situational prioritization of information. Extending the work on conflicts as negative signals, the findings indicate that when prioritization fails to resolve conflict, negative judgments of conflicting stimuli are easier than positive judgments as indicated by decision times and engagement of dACC and DLPFC.

Supplementary data

Supplementary data are available at *SCAN* online.

Notes.

1. Given the traits and situations included in the study, a majority of the judgments related to the competence of the target person in a given situation (Supplementary Material S2).
2. Participants additionally rated how positively (negatively) they evaluated each target persons' most positive (negative) traits on a scale ranging from not at all (0) to very (100). Ratings were combined in line with Thompson Zanna, and Griffin (1995): Ambivalence = (Pos + Neg)/2 −|Pos − Neg|. This measure has been moved to the notes due to word constraints and because we expected the same pattern of results as on the subjective ambivalence scale. Results on this measure indeed confirmed the other self-report results that targets described by mixed traits elicited more ambivalence ($M = 48.05$, SE = 5.14) than did targets described by consistent traits [$M = -30.21$, SE = 4.61; $t(19) = -9.405$, $P < 0.001$, $r = 0.91$].
3. For exploratory reasons, participants also responded to the same questions presented in the scanner on a continuous scale ranging from definitely not (0) to definitely yes (100). This measure was not analyzed for this study.
4. We verified the results of the MLM by running an ANOVA. Note that because of the unbalanced design, ANOVA is

inferior to MLM in this case. This analysis resulted in similar findings; the details can be found in the Supplementary Material S4.

5. Activation in right DLPFC explained response behavior [$\chi^2(1) = 4.47$, $P = 0.03$; left DLPFC: $\chi^2(1) = 6.50$, $P = 0.01$] next to decision times [$\chi^2(1) = 10.75$, $P = 0.001$; left DLPFC: $\chi^2(1) = 9.48$, $P = 0.002$]. Whereas there was no main effect of trial type on the valence of the judgments participants made, we observed that the effect of decision time on the direction of judgments depended on trial type [$\chi^2(2) = 16.69$, $P = 0.0002$; left DLPFC: $\chi^2(2) = 14.89$, $P = 0.0006$].
6. We observed that the direction of response behavior could be explained by (participant- and valence-centered) activation in dACC [$\chi^2(1) = 3.70$, $P = 0.05$] next to (participant- and valence-centered) decision times [$\chi^2(1) = 9.35$, $P = 0.002$]. Variation in decision times also predicted the valence of judgments dependent on trial type [$\chi^2(2) = 15.21$, $P = 0.0005$].

Acknowledgments

We would like to thank Eveline Crone and Vincent Man for their helpful comments on earlier versions of this manuscript, and Anthony Romyn for his contribution to the data analysis.

References

Alexander, W.H., Brown, J.W. (2015). Hierarchical error representation: a computational model of anterior cingulate and dorsolateral prefrontal cortex. *Neural Computation*, **27**, 2354–410 http://doi.org/10.1162/NECO.

Badre, D., Wagner, A.D. (2004). Selection, integration, and conflict monitoring: assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, **41**(3), 473–87 http://doi.org/10.1016/S0896-6273(03)00851-1.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**(1), 1–48 http://doi.org/10.18637/jss.v067.i01.

Bhanji, J.P., Beer, J.S. (2013). Dissociable neural modulation underlying lasting first impressions, changing your mind for the better, and changing it for the worse. *The Journal of Neuroscience*, **33**(22), 9337–9344 https://doi.org/10.1523/JNEUROSCI.5634-12.2013.

Boorman, E.D., Rushworth, M.F., Behrens, T.E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, **33**(6), 2242–53 http://doi.org/10.1523/JNEUROSCI.3022-12.2013.

Botvinick, M.M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, **7**(4), 356–66 http://doi.org/10.3758/CABN.7.4.356.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**(3), 624–52 http://doi.org/10.1037/0033-295X.108.3.624.

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), 137–49 http://doi.org/10.1162/jocn.

Cacioppo, J.T., Gardner, W.L., Berntson, G.G. (1997). Beyond bipolar conceptualizations and measures: the case of attitudes and evaluative space. *Personality and Social Psychology Review*, **1**(1), 3–25.

Carter, C.S., van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, **7**(4), 367–79 http://doi.org/10.3758/cabn.7.4.367.

Cunningham, W.A., Zelazo, P.D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends in Cognitive Sciences*, **11**(3), 97–104 http://doi.org/10.1016/j.tics.2006.12.005.

Cunningham, W.A., Zelazo, P.D., Packer, D.J., Van Bavel, J.J. (2007). The iterative reprocessing model: a multilevel framework for attitudes and evaluation. *Social Cognition*, **25**(5), 736–60 http://doi.org/10.1521/soco.2007.25.5.736.

Danziger, S., Levav, J., Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, **108**(17), 6889–92 http://doi.org/10.1073/pnas.1018033108.

Dreisbach, G., Fischer, R. (2012). Conflicts as aversive signals. *Brain and Cognition*, **78**(2), 94–8 http://doi.org/10.1016/J.BANDC.2011.12.003.

Ebitz, R.B., Hayden, B.Y. (2016). Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. *Nature Neuroscience*, **19**(10), 1278–9 http://doi.org/10.1038/nn.4387.

Fox, J., Weisberg, S. (2019). *An {R} Companion to Applied Regression*, 3rd edn, Thousand Oaks, CA: Sage, Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Fritz, J., Dreisbach, G. (2013). Conflicts as aversive signals: conflict priming increases negative judgments for neutral stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, **13**(2), 311–7 http://doi.org/10.3758/s13415-012-0147-1.

Gray, J.A., McNaughton, N. (2000). *The Neuropsychology of Anxiety: An Enquiry Into the Functions of Septohippocampal Theories*, 2nd edn, Oxford, UK: Oxford University Press, http://doi.org/10.1017/S0140525X00013170.

Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(44), 18120–5 http://doi.org/10.1073/pnas.1109322108.

Heilbronner, S.R., Hayden, B.Y. (2016). Dorsal anterior cingulate cortex: a bottom-up view. *Annual Review of Neuroscience*, **39**(1), 149–70 http://doi.org/10.1146/annurev-neuro-070815-013952.

Hughes, B.L., Zaki, J., Ambady, N. (2017). Motivation alters impression formation and related neural systems. *Social Cognitive and Affective Neuroscience*, **12**(1), 49–60 http://doi.org/10.1093/scan/nsw147.

Inzlicht, M., Bartholow, B.D., Hirsh, J.B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, **19**(3), 126–32 http://doi.org/10.1016/j.tics.2015.01.004.

Jenkinson, M., Bannister, P., Brady, M., Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, **17**(2), 825–41 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12377157.

Jenkinson, M., Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, **5**(2), 143–56 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11516708.

Kim, H., Somerville, L.H., Johnstone, T., Alexander, A.L., Whalen, P.J. (2003). Inverse amygdala and medial prefrontal cortex responses to surprised faces. *NeuroReport*, **14**(18), 2317–22 http://doi.org/10.1097/01.wnr.0000101520.44335.20.

Kim, H., Somerville, L.H., Johnstone, T., *et al.* (2004). Contextual modulation of amygdala responsivity to surprised faces.

Journal of Cognitive Neuroscience, **16**(10), 1730–45 http://doi.org/10.1162/0898929042947865.

Kolling, N., Wittmann, M.K., Behrens, T.E.J., Boorman, E.D., Mars, R.B., Rushworth, M.F.S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience*, **19**(10), 1280–5 http://doi.org/10.1038/nn.4382.

Kool, W., Shenhav, A., Botvinick, M.M. (2017). Cognitive control as cost–benefit decision making. In: Egner, T., editor. *The Wiley Handbook of Cognitive Control*, 1st edn, Chichester, West Sussex, UK: John Wiley & Sons, Ltd., pp. 167–89.

Kriegeskorte, N., Lindquist, M.A., Nichols, T.E., Poldrack, R.A., Vul, E. (2010). Everything you never wanted to know about circular analysis, but were afraid to ask. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, **30**(9), 1551–7 http://doi.org/10.1038/jcbfm.2010.86.

Lieberman, M.D., Cunningham, W.A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience*, **4**(4), 423–8 http://doi.org/10.1093/scan/nsp052.

Luttrell, A., Stillman, P.E., Hasinski, A.E., Cunningham, W.A. (2016). Neural dissociations in attitude strength: distinct regions of cingulate cortex track ambivalence and certainty. *Journal of Experimental Psychology: General*, **145**(4), 419–33 http://doi.org/10.1037/xge0000141.

Ma, N., Vandekerckhove, M., Baetens, K., Van Overwalle, F., Seurinck, R., Fias, W. (2012). Inconsistencies in spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, **7**(8), 937–50 http://doi.org/10.1093/scan/nsr064.

Ma, N., Vandekerckhove, M., van Overwalle, F., Seurinck, R., Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: spontaneous inferences activate only its core areas. *Social Neuroscience*, **6**(2), 123–38 http://doi.org/10.1080/17470919.2010.485884.

Mende-Siedlecki, P., Cai, Y., Todorov, A. (2013). The neural dynamics of updating person impressions. *Social Cognitive and Affective Neuroscience*, **8**(6), 623–31 http://doi.org/10.1093/scan/nss040.

Miller, E.K., Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, **24**, 167–202.

Mumford, J.A., Turner, B.O., Ashby, F.G., Poldrack, R.A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, **59**(3), 2636–43 http://doi.org/10.1016/j.neuroimage.2011.08.076.

Neta, M., Norris, C.J., Whalen, P.J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion (Washington, D.C.)*, **9**(5), 640–8 http://doi.org/10.1037/a0016819.

Neta, M., Tong, T.T. (2016). Don't like what you see? Give it time: longer reaction times associated with increased positive affect. *Emotion*, http://doi.org/10.1037/emo0000181.

Nohlen, H.U., van Harreveld, F., Rotteveel, M., Barends, A.J., Larsen, J.T. (2016). Affective responses to ambivalence are context-dependent: a facial EMG study on the role of inconsistency and evaluative context in shaping affective responses to ambivalence. *Journal of Experimental Social Psychology*, **65**, 42–51 http://doi.org/10.1016/j.jesp.2016.02.001.

Nohlen, H.U., van Harreveld, F., Rotteveel, M., Lelieveld, G.-J., Crone, E.A. (2014). Evaluating ambivalence: social-cognitive and affective brain regions associated with ambivalent decision-making. *Social Cognitive and Affective Neuroscience*, **9**(7), 924–31 http://doi.org/10.1093/scan/nst074.

Ohman, A., Flykt, A., Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology. General*, **130**(3), 466–78 Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11561921.

Priester, J.R., Petty, R.E. (1996). The gradual threshold model of ambivalence: relating the positive and negative bases of attitudes to subjective ambivalence. *Journal of Personality and Social Psychology*, **71**(3), 431–49 http://doi.org/10.1037/0022-3514.71.3.431.

R Core Team (2013). R Development Core Team. *R: A Language and Environment for Statistical Computing*, **55**, 275–86.

Rozin, P., Royzman, E.B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, **5**(4), 296–320 http://doi.org/10.1207/S15327957PSPR0504_2.

Rushworth, M.F.S., Noonan, M.A.P., Boorman, E.D., Walton, M.E., Behrens, T.E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, **70**(6), 1054–69 http://doi.org/10.1016/j.neuron.2011.05.014.

Sallet, J., Mars, R.B., Quilodran, R., Procyk, E., Petrides, M., Rushworth, M.F.S. (2011). Neuroanatomical basis of motivational and cognitive control: a focus on the medial and lateral prefrontal cortex. In: Mars, R.B., Sallet, M.F.S., Rushwort, J., Yeung, N., editors. *Neural Basis of Motivational and Cognitive Control*, Cambridge, MA: MIT Press, pp. 5–20 http://doi.org/10.7551/mitpress/9780262016438.001.0001.

Schouppe, N., Braem, S., De Houwer, J., *et al.* (2015). No pain, no gain: the affective valence of congruency conditions changes following a successful response. *Cognitive, Affective, & Behavioral Neuroscience*, **15**(1), 251–61 http://doi.org/10.3758/s13415-014-0318-3.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Andrew, S., Winter, J.J., Davidson, R.J. (2011). The integration of negative affect, pain, and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, **12**(3), 154–67 http://doi.org/10.1038/nrn2994.The.

Shenhav, A., Botvinick, M., Cohen, J. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, **79**(2), 217–40 http://doi.org/10.1016/j.neuron.2013.07.007.

Smith, S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, **17**(3), 143–55 http://doi.org/10.1002/hbm.10062.

Snyder, A.I., Tormala, Z.L. (2017). Valence asymmetries in attitude ambivalence. *Journal of Personality and Social Psychology*, **112**(4), 555–76 http://doi.org/10.1037/pspa0000075.

Somerville, L.H., Heatherton, T.F., Kelley, W.M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, **9**(8), 1007–8 http://doi.org/10.1038/nn1728.

Stillman, P.E., Van Bavel, J.J., Cunningham, W.A. (2015). Valence asymmetries in the human amygdala: task relevance modulates amygdala responses to positive more than negative affective cues. *Journal of Cognitive Neuroscience*, **27**(4), 842–51 http://doi.org/10.1162/jocn_a_00756.

Thompson, M.M., Zanna, M.P., Griffin, D.W. (1995). Let's Not Be Indifferent About (Attitudinal) Ambivalence. In R. E. Petty & J. A. Krosnick (Eds.) *Attitude strength: Antecedents and consequences*, (pp. 361–386) Hillsdale, NJ: Erlbaum.

Tottenham, N., Phuong, J., Flannery, J., Gabard-Durnam, L., Goff, B. (2013). A negativity bias for ambiguous facial-expression valence during childhood: converging evidence from behavior and facial corrugator muscle responses. *Emotion*, **13**(1), 92–103 http://doi.org/10.1037/a0029431.