Rapid approach-avoidance responses to emotional displays reflect value-based decisions: Neural evidence from an EEG study

Rocco Mennella, Emma Vilarem, Julie Grezes

NeuroImage 222 (2020) 117253

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Rapid approach-avoidance responses to emotional displays reflect value-based decisions: Neural evidence from an EEG study

Rocco Mennella*, Emma Vilarem, Julie Grèzes*

Cognitive and Computational Neuroscience Laboratory (LNC²), Inserm U960, Department of Cognitive Studies, École Normale Supérieure, PSL University, 29 rue d'Ulm, 75005 Paris, France

ARTICLE INFO

Keywords:
Emotional displays
Approach-avoidance
Goal-directed
Value encoding
Diffusion models
EEG

ABSTRACT

The ability to swiftly and accurately respond to others' non-verbal signals, such as their emotional expressions, constitutes one of the building blocks for social adaptation. It is debated whether rapid action tendencies to socio-emotional signals solely depend upon stimulus-evoked pre-decisional motor bias or can also engage goal-directed (decisional) processes that involve the arbitration between action alternatives. Here, we used drift diffusion models (DDMs) of choice and electroencephalography (EEG) to investigate the impact of threat-signaling individuals (angry or fearful) on spontaneous approach-avoidance decisions. Participants choose to avoid angry individuals more often than fearful ones and this effect was stronger for intense expressions. Diffusion models showed that this pattern of choice was accounted for by a process of value-based evidence accumulation, suggesting an active competition between action options. At the brain level, we found that EEG activity preceding movement initiation (200 ms) in a mid-frontal cluster of electrodes – sourced in the orbital and ventromedial frontal cortices – encoded value difference between chosen and unchosen options, thus predicting participant's choices on a trial-by-trial basis. Furthermore, value difference also modulated EEG signal during feedback about the decision. Altogether, the present findings convincingly support the underestimated influence of implicit goal-directed mechanisms in approach-avoidance responses to socio-emotional signals.

1. Introduction

Emotional facial expressions inform others about the affective states and potential behavioral intentions of the emitter (Waller et al., 2017), conveying action demands to the perceiver (Horstmann, 2003). Accordingly, perceiving emotional expressions has a direct influence on the observer's behavior (Dezecache et al., 2015) generating, for instance, action tendencies to approach or avoid (Hammer and Marsh, 2015; Marsh et al., 2005). Such action tendencies to emotional displays are ultimately thought to facilitate social interactions and adaptive behavior (Keltner and Haidt, 1999). Despite their importance, the neuro-cognitive mechanisms underlying the generation of approach-avoidance tendencies to emotional displays are not fully understood, being currently a subject of debate (Bach and Dayan, 2017).

One the one hand, action tendencies may consist in the direct activation of a response representation (e.g., avoidance) by some features of the perceived emotional display (e.g., eye frowning of an angry face), due to a pre-existing stimulus-response association (e.g., Frijda, 1986; Lang et al., 1990; Öhman, 1986). In this scenario, the perceived emotional display would automatically strengthen the motor representation of the action that leads to threat avoidance and this stimulus-evoked pre-decisional bias would influence the final action choice. Although they can be overridden or refined through top-down control mechanisms, pre-decisional motor biases are nonetheless expected to have an influence on people's responses to emotional displays (e.g., Bramson et al., 2018; Roelofs et al., 2009). Such biases undeniably promote survival, notably in high threat situations but, as such, they might contribute only to a small degree to the action repertoire adopted in realistic environments (Cain, 2019).

On the other hand, converging research on defensive behavior in animals (Evans et al., 2019; LeDoux and Daw, 2018) and on approach-avoidance in humans (Eder and Hommel, 2013; Moors et al., 2019, 2017; Rotteveel and Phaf, 2004; Schlund et al., 2016) suggested that action tendencies might not always be automatically elicited by emotional stimuli in a pre-decisional manner, but that they can be the result of goal-directed processes. Goal-directed behaviors are said to be emitted, as opposed to be elicited or triggered, since they are actively selected on the basis of the value assigned to learned action-outcome contingencies (LeDoux and Daw, 2018). Interestingly, goal-directed processes subtending approach-avoidance tendencies to emotion are suggested to

* Corresponding authors.
E-mail addresses: rocco.mennella@gmail.com (R. Mennella), julie.grezes@ens.fr (J. Grèzes).

https://doi.org/10.1016/j.neuroimage.2020.117253
Received 22 January 2020; Received in revised form 3 July 2020; Accepted 9 August 2020
Available online 13 August 2020

often take place very rapidly outside consciousness in both human and nonhuman animals (LeDoux and Daw, 2018). Here, we use the term "decisional" process to describe this rapid arbitration between action alternatives, which leads to goal-directed behaviors, in opposition to "pre-decisional motor biases", which precede arbitration between alternatives and underlie stimulus-evoked behaviors.

Important insights on the neurocognitive processes subtending approach-avoidance behaviors to emotional stimuli in healthy humans (e.g., Chen and Bargh, 1999; De Houwer et al., 2001; Solarz, 1960) and in clinical populations (e.g., Jones et al., 2013; Taylor and Amir, 2012; Wiers et al., 2011) have been brought so far by studies employing classical paradigms, such as the Approach Avoidance Tasks (AAT) and the Manikin Task. Nonetheless, the forced-choice nature of these tasks intrinsically limits the influence of decisional processes, which select the best option among action alternatives on the basis of the value of their expected outcomes. To reveal both pre-decisional and decisional components of behavioral responses to emotional stimuli, approach-avoidance paradigms involving spontaneous decision-making are needed (Paré and Quirk, 2017). Similarly to what happens in everyday life, this implies requiring participants to choose among different options for action (Moors et al., 2017). We therefore developed a new paradigm to assess the impact of task-irrelevant threat-signaling expressions (anger and fear) on action decisions between two alternative targets for action (Vilarem et al., 2019). This task revealed that the presence of emotional individuals influenced spontaneous approach-avoidance responses with anger, but not fear, being clearly associated with behavioral avoidance.

In agreement with the idea that approach-avoidance behaviors are subtended by a rapid value-based decision, we originally suggested that the presence of task-irrelevant threat-signaling expressions (anger and fear) might implicitly change the value of each available action option (Vilarem et al., 2019). In particular, anger, a direct signal of aggression, would increase the expected value of the action leading to avoidance. We thus hypothesized that avoiding an angry individual is a desirable outcome for most subjects, who implicitly choose it over its alternative.

In the present study, we aimed at testing this hypothesis more directly, elucidating the neuro-cognitive mechanisms underlying the observed pattern of behavioral results. Firstly, it has recently been argued that computational models of decision offer the opportunity to formalize classic theoretical concepts from affective sciences into quantitative mathematical parameters (Roberts and Hutcherson, 2019). For instance, in simple two-choice tasks, Drift Diffusion Models (DDMs) (Ratcliff and McKoon, 2008) allow translating the raw proportion of choice and the distribution of response times into parameters representing the contribution of different cognitive processes. In our task, the observed behavior may have emerged from a modulation of (1) the pre-decisional bias, i.e. the starting point of the accumulation process, which would reduce the amount of stimulus evidence needed for producing the response (e.g., avoidance); (2) the decisional process itself, i.e. the rate of evidence accumulation; or (3) both. Following the interpretation proposed in Vilarem et al. (2019), we expected subjects' behavior in the presence of emotional displays to be better accounted for by changes in the decisional process, i.e. increased rate of evidence accumulation, rather than by a shift in pre-decisional bias toward avoidance responses.

Secondly, to further test this hypothesis, we used the high temporal resolution of electroencephalography (EEG) to assess whether the difference in the value of the action alternatives is represented in the brain. Accordingly, such neural encoding of value differences would be expected in case the tendency to approach/avoid follows a value-based decision. More precisely, we expected to observe an early neural encoding of the difference in value between chosen vs. unchosen action options in brain regions known to correlate with option values (Bartra et al., 2013), between stimulus presentation and movement initiation (Hunt et al., 2015, 2013, 2012), as well as after providing feedback about the decision (McCoy et al., 2003).

2. Material and methods

2.1. Participants

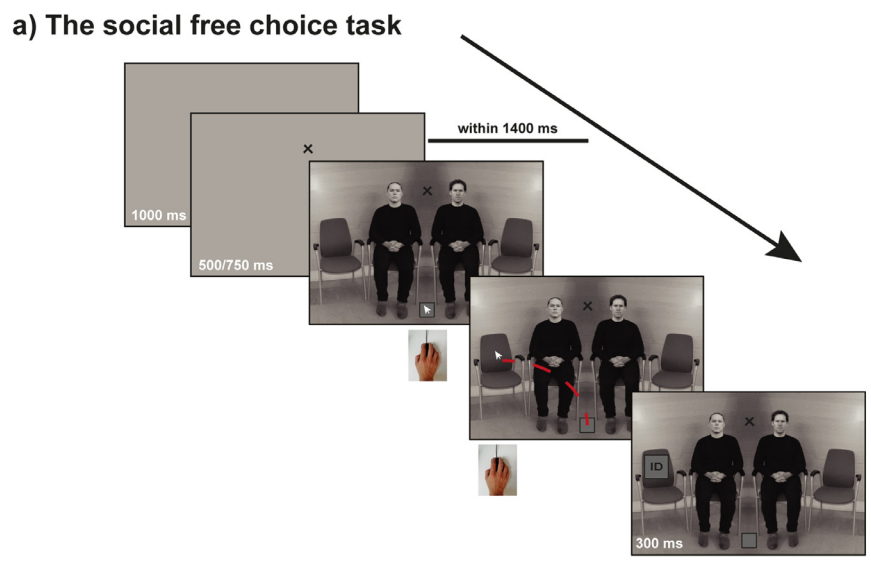
As we consistently replicated the main effect of Emotion (anger vs. fear) in our previous studies (Vilarem et al., 2019), we calculated the a-priori of required sample size for detecting a significant effect of Emotion on the proportion of choice for the present experiment (alpha = 0.05; power = 0.80). We used the raw data of the study presented in the supplementary material of Vilarem and colleagues' paper, which had the larger sample size, i.e. 40 subjects. We recalculated the ANOVA on the proportion of away responses, obtaining an $\eta^2_p = 0.268$ for the effect of Emotion, corresponding to a required sample size of 25 subjects. We aimed at enrolling 32 participants into the study, to compensate for potential withdrawal and technical problems. One subject never showed up, four were excluded due to noisy EEG data and one due to a technical failure in the synchronization of stimulus markers. The final sample (N = 26, 14 females) had a mean age of 23.7±3.4 (age range = 19–34).

All participants involved in the study reported no history of neurological or psychiatric disorders. The experimental protocol was approved by INSERM and the local research ethics committee (Comité de protection des personnes Ile de France III – Project CO7-28, N° Eudract: 207-A01125-48), and was carried out in accordance with the Declaration of Helsinki. The participants provided informed written consent and were compensated for their participation.

2.2. The social free-choice task

The social free-choice task (Fig. 1a) was adapted from a previous study from our lab (Vilarem et al., 2019). Subjects were presented with a scene representing a waiting room with four seats, where the two middle seats were occupied by two individuals (a pair of females or males) and the outer seats were empty. Each scene was the composite of one template female or male hemi-scene (photograph depicting either one female or one male sitting next to an empty seat; 835 × 1050 pixels) juxtaposed to its mirrored version, on which the faces were superimposed. Faces were selected from the RadBoud Faces Database (Langner et al., 2010) before being adapted to the template female or male body. One actor of the pair always displayed a neutral expression, while the other displayed either a neutral, angry or a fearful expression. We used ten (five males, five females) fixed pairs of identities matched for facial trustworthiness and threat traits (for details see Vilarem et al., 2019). Faces varied in emotion (neutral, angry or fearful expressions) and in intensity (4 levels of morphs for anger and fear were created from the neutral to the emotional expression using a simple linear morphing transformation). Each morph level was equalized between anger and fear in perceived emotional intensities (for details see El Zein et al., 2015).

At the beginning of each trial, a grey screen appeared for 1000 ms, then a fixation cross was superimposed upon the grey screen for a time varying between 500 and 750 ms. After the fixation, the scene (luminance centered on the mean across all the images; on Matlab, mean= 0.428, sd = 0.202) appeared and remained on the screen until a correct response was registered, or until a maximum time of 1400 ms in the case of no response. Participants were asked to indicate the seat they would like to occupy, maintaining fixation on the cross displayed between the faces throughout the trial. In order to make their choice, participants had to left-click on the mouse, move the cursor from the bottom center of the scene and release the click on the chosen seat. The cursor was automatically re-centered at every new trial. Participants were required to make spontaneous choices and were informed that there were no correct or incorrect responses. Nevertheless, their movements needed to be properly performed for their responses to be registered. A proper movement was defined by the release of the click on one of the seats within 1400 ms after scene onset. After a response, a picture of the participant's face – taken prior to the beginning of the experiment – was superimposed on the scene at the release location for 300 ms,



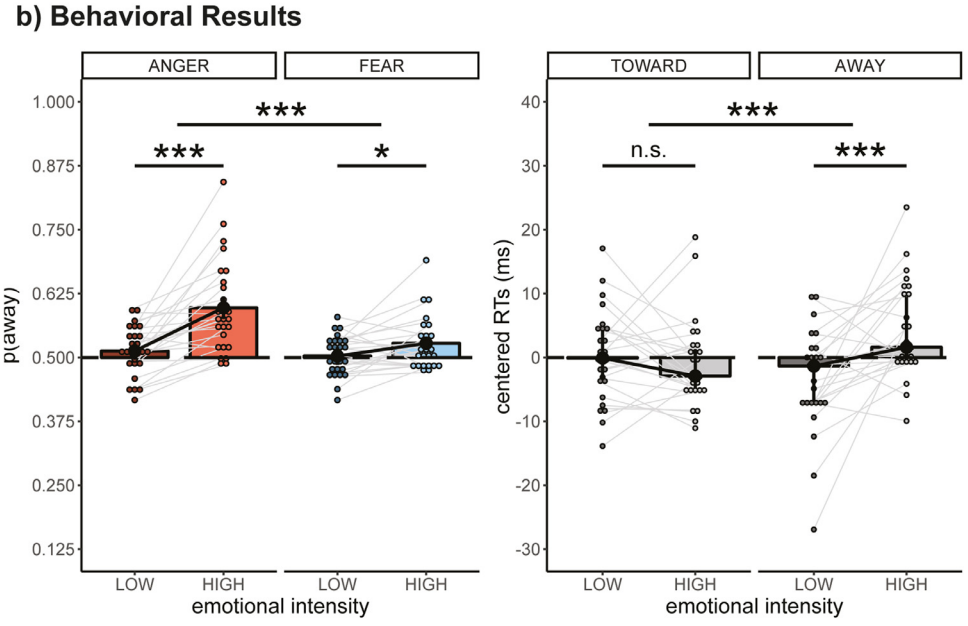


Fig. 1. Task and behavioral results. (a) The social free-choice task. Time course of a trial where participants are asked to indicate where they would like to sit. The identities displayed were not used in the real experiment and were selected for illustration purposes only, following the guidelines of the Redboud Faces Database. (b) Behavioral results. Left plot: black dots and vertical lines represent the mean and the within-subject confidence interval of the mean proportion of away responses for each subject (smaller colored dots). Grey lines connect subjects' means across conditions. Right plot: black dots and vertical lines represent the median and the confidence interval of the median RTs for each subject, centered on the subject mean to get rid of between-subject variability (smaller colored dots). Grey lines connect subjects' medians across conditions. $^{***} = p < .001$; $^{**} = p < .01$; $^{*} = p < .05$; n.s. $= p > .05$.

providing a feedback that the choice had been correctly registered. All factors were fully counterbalanced and pseudorandomized across participants. This resulted in 480 trials – 10 actor pairs × (2 emotional expressions × 4 levels of morphs + 1 neutral expression × 4 repetitions) × 2 emotional actor's identity × 2 emotional actor's side – which we doubled to increase the number of trials for computational modelling. Thus, the entire experiment consisted in 960 trials, subdivided into ten blocks of 96 trials. The experiment was run using Psychtoolbox on Matlab R2012b.

2.3. Procedure

Upon arrival at the laboratory, the participants read and signed an informed consent form. Participants were then seated in a dimly lit, sound-attenuated room and EEG electrodes were attached. The distance from eye to screen was 60 cm, so that the eccentricity to the central fixation cross was of 4.5° for the center of the faces, and of 8° for the center of the seats. Participants were given the instructions for the social free-choice task, and were also instructed to avoid blinking during picture presentation and to maintain fixation. Before completing the experiment, all the subjects went through training blocks depicting only neutral individuals to practice with the task and the response required, until accuracy reached at least 60%. Both during the training and the task, they were informed of their percentage of correct executions at the end of each block and were asked to maximize it.

2.4. EEG recording and processing

Using a BioSemi headcap with active electrodes, the EEG was continuously recorded from 64 scalp sites, with CMS/DRL reference electrodes. The EEG signal was amplified using an ActiveTwo AD-box amplifier (BioSemi), low-pass filtered online (250 Hz) and digitized at 1000 Hz. Pre-processing of the EEG signal was run in EEGlab (Delorme and Makeig, 2004). The signal was referenced offline to an average reference, down-sampled at 512 Hz, band-pass filtered between 1 and 32 Hz and epoched from 2 s before to 3 s after stimulus' onset. Epochs were visually inspected and discarded if containing muscular artifacts and noisy electrodes were interpolated averaging the adjacent electrodes. Finally, blink artifacts were corrected through manual removal of the corresponding ICA components. The EEG signal was further epoched between −1.5 s and 2 s around stimulus' appearance and between −2 to 1.2 s around release time (i.e., when the cursor was released on the

chosen chair), and further down-sampled to 250 Hz, to facilitate data handling for statistical analyses.

Source analysis was performed using Brainstorm (Tadel et al., 2011). Full noise covariance matrices were computed using the −0.5 to −0.1 time-window before stimulus' onset. A source space consisting of 7001 dipoles constrained to the cortical mantle of the standard ICBM152 template brain provided in Brainstorm was chosen, and the forward model was calculated using a 3D overlapping sphere method. Kernel inversion matrices (7001 vertices * 64 electrodes) were computed for each subject, based on a "depth-weighted" linear L2-minimum norm estimate, using Brainstorm default settings (depth weighting (Order[0,1]) = 0.5, Maximal amount = 10; Noise covariance regularization = 0.1; $1/\lambda = 3$). Subjects' kernels were then extracted and multiplied to each trial to perform single trial regressions at the source level.

2.5. Statistical analyses: behavioral data

2.5.1. Linear models

Behavioral data analysis was run on the same epochs accepted for the final EEG analysis, to allow EEG-behavior correlations. EEG accepted trials were filtered within each subject below the 1st and above the 99th percentile of both click time (i.e., time between stimulus' appearance and mouse click) and movement time (i.e., time between mouse click and mouse release on the chosen chair), to eliminate response anticipations and atypical movements.

Median click times (from now on "Reaction Times; RTs") and movement times for each subject were calculated and then log-transformed to normalize their distribution. Responses were coded as follows: if the subject sat on the chair far from the individual displaying threat-related facial expressions, the response was coded "away". On the other hand, if the subject sat on the chair close to the individual displaying threat-related facial expressions, the response was coded "toward". To facilitate the fitting of our DDM models, the Intensity factor was re-coded on 2 levels (i.e., low-intensity = level1+ level2 and high-intensity = level3+ level4), obtaining a minimum number of 111 trials per subject for each condition of the interaction between Emotion and Intensity.

First, repeated-measures ANOVAs were fitted on the variables of interest using the "EzANOVA" function of the "Ez" package in R (Lawrence, 2016). For the proportion of choice, a repeated-measures ANOVA on the mean proportion of away responses was employed, with Emotion (Anger, Fear) and Intensity (High, Low) as within-subject factors.

For both RTs and movement times, repeated-measures ANOVAs on the log-transformed median scores for each subject were fitted, with Emotion, Intensity and Side (Away, Toward) as within-subject factors.

In order to substantiate our behavioral results, we also fitted generalized mixed effects models on full-trial database, with the same fixed effects as the ANOVAs, allowing a random intercept to each subject ("EzMixed" function of the "Ez" package). A binomial distribution was assumed to predict away (1) vs. toward (0) responses at each trial, and a Gaussian distribution was assumed to predict log-transformed RTs. Assessment of each effect of interest was inferred by comparing a model that contained the effect of interest plus any lower order effects with a "restricted" model with only the lower order effects. Comparison was made through a likelihood ratio, corrected for the different complexity of the two models (Glover and Dixon, 2004). To correct the likelihood, we choose the Akaike's Information Criterion (AIC), which in the context of mixed effects models has been demonstrated to be asymptotically equivalent to cross-validation (Fang, 2011). The complexity-corrected likelihood ratio was converted in the log-base-2 scale, so that resulting values can be discussed as representing "bits of evidence" for or against the evaluated effect (Lawrence, 2016).

2.5.2. Drift diffusion models

We employed the fast-dm software (version 30.2) (Voss et al., 2015; Voss and Voss, 2008, 2007) to fit DDMs on the subjects' choices and RTs. DDMs are used to infer the cognitive processes involved in binary decision tasks: the decision process itself, defined as the rate (v) at which evidence for one of the choices is accumulated, the threshold (a) that represents the amount of information which separates the two alternative choices, the pre-decisional bias which is mapped on the starting point (z) – the closer the starting point to a threshold, the less information is needed to decide for that option, and finally the non-decision time ($t0$) that captures stimulus encoding and response execution, which respectively precede and follow the decision process (Voss et al., 2015).

In DDMs the bias parameter z is typically thought to represent a pre-decisional preference for one of the two action options, for instance to respond with our left or right hand and this irrespective of the stimulus. Here, we consider a slightly different interpretation of the pre-decisional bias, which takes into considerations knowledge from affective neurosciences (Roberts and Hutcherson, 2019). Indeed, it is suggested that the process of highly salient stimuli, such threatening ones, takes place very rapidly (LeDoux, 2012, 1996). Automatic tendencies to approach or avoid emotional stimuli are suggested to be sustained by amygdala's direct connections to subcortical (e.g., Hashemi et al., 2019) and cortical motor centers (e.g., Grèzes et al., 2014; Toschi et al., 2017). Here, we propose that, in everyday situation, where the movements that will allow us to approach or avoid are not predetermined, the praecox identification of the source of the salient stimuli can result in a rapid action disposition (pre-decisional motor bias) toward approach or avoidance, captured by the z parameter of the model.

On the base of this assumption, four models were run to disambiguate between the pre-decisional and decisional hypotheses: (1) a null model, where none of the parameters varied as a function of our factors of interest (Emotion and Intensity); (2) a model where only the starting point (z) was allowed to vary as a function of our factors of interest; (3) a model where only the drift rate (v) was allowed to vary as a function of our factors of interest; (4) a model where both z and v were allowed to vary as a function of our factors of interest. Due to the moderate number of trials per condition, we simplified the models as much as possible, in several ways: the thresholds of the model were associated with away (upper threshold) and toward responses (lower threshold); furthermore, the inter-trial variabilities of drift rate and threshold (but not of the non-decision time) were fixed to zero, since a proper fit of these parameters is particularly challenging, especially with small to medium-sized trials number (Lerche and Voss, 2016).

The Minimum-Norm optimization was used for parameter's estimation. The AIC was extracted for each subject and the mean AIC for each model was computed to allow model comparison. In addition to the mean AIC, which could be affected by subjects' heterogeneity, we relied on a hierarchical Bayesian model selection criterion, in which models are random variables (Rigoux et al., 2014; Stephan et al., 2009). The model estimates the parameters of a Dirichlet distribution which describes models' probabilities, which in turn define a multinomial distribution over model space, allowing to calculate the exceedance probability of the winning model being more likely than the others. Finally, parameter estimates from the winning drift diffusion model were tested using repeated-measures ANOVAs.

2.6. Statistical analyses: EEG data

For EEG analyses, we used a General Linear Models approach (GLM), previously validated in our laboratory (El Zein et al., 2015; Patron et al., 2019; Wyart et al., 2015, 2012). In brief, this method consists in fitting single-trial regressions models in order to evaluate the encoding of experimental factors in the EEG signal, at each electrode and time point, for each subject:

$$EEG \sim factors\ of\ interest$$

Such GLMs produce an electrode by time matrix of beta values for each subject, which represents the strength of the linear relation between predictors and dependent variables. All regression-based analyses

of the EEG data were followed by a second-level analysis at the group level, to assess the significance of the observed effects across participants. In order to perform second-order statistics, we choose a cluster-based approach (Maris and Oostenveld, 2007) to control over the type I error rate arising from multiple comparisons across electrodes and time points. First, standard parametric tests (*t*-tests against zero) were run across electrodes and time points. The resulting values were thresholded (p_{thresh} = .05), and the pairing between experimental conditions and EEG signals was shuffled pseudo-randomly 2000 times. The maximal cluster-level statistics (e.g., the sum of *t*-values across contiguously significant time points at the threshold level) were extracted for each shuffle to compute a 'null' distribution of effect sizes. For each significant cluster in the original (non-shuffled) data, we computed the proportion of clusters in the null distribution whose statistics exceeded the one obtained for the cluster in question, corresponding to its cluster-corrected *p*-value. Clusters with a $p_{corr} < .05$ were considered significant.

For each subject, in order to control for the influence of movement on EEG activity, we first calculated the residuals of the EEG activity predicted by the movement parameters (i.e., RTs and movement times), log-transformed to normalize their distribution and z-scored across conditions:

$$EEG \sim zscore(\log(RT)) + zscore(\log(Movement\ Time))$$

at each electrode and time point (or at each voxel and time point for EEG sources). Subsequent GLMs were run on the residuals of this model.

In order to test whether the difference in value between the chosen vs. the unchosen action options was represented in the brain, we built a "Value Difference" regressor based on the intensity of the emotion displayed, its side on the scene and the response of the subject. To do so, the seat close to the threatening individual was considered as the punishing option, which would have a negative value proportionate to the level of emotional intensity of the threatening expression (coded as 0.5 = level 1, 1.5 = level 2, 2.5 = level 3, 3.5 = level 4). The other seat, close to the neutral individual, would have a non-negative value (i.e., zero). Therefore, if for instance the subject decided to sit far from an individual expressing anger at level 3 (i.e. away response), the difference in value for her choice would correspond to Chosen option – Unchosen Option = 0 – (–2.5) = +2.5, therefore a positive value difference. On the contrary, if in the same trial she chose to sit close to the individual expressing anger (i.e. toward response), the difference in value for her choice would correspond to Chosen option – Unchosen Option = –2.5 – 0 = –2.5, therefore a negative value difference. Such calculation directly follows our hypothesis that emotion has a role in changing the respective value of competing action plans (Vilarem et al., 2019).

Separately for anger and fear, we ran the following GLM on stimulus-locked data (i.e., residuals) for each subject:

$$EEG \sim Value\ Difference(-3.5\ to\ 3.5)$$

For stimulus-locked analysis, we selected a time window ranging from stimulus appearance to the longest of the median RTs across subjects (i.e., from 0 to 580 ms). For feedback-locked analysis, we selected a time window ranging from the longest median movement time (i.e., 680 ms) to the end of the response-related feedback (i.e., 300 ms) (total range = −680 to 300 ms).

Finally, the same models were run on each voxel's residual source activity, after controlling for movement parameters. Regression parameters at the source levels were smoothed (6 mm) and then contrasted to zero with a *t*-test.

2.7. Statistical analyses: EEG-behavior

We further tested whether the EEG encoding of value differences were related to the quality of evidence accumulation, across subjects. EEG activity in the significant clusters for both stimulus- and feedback-locked analyses was therefore correlated with difference in drift rate between high and low intensity trials. Moreover, since the duration of the non-decision phase could have an impact on the availability of time necessary for comparing the value of the alternatives, we also tested the correlation between t0 and the EEG encoding of value.

2.8. Data and code availability

Due to constraints of our ethical agreement, data and custom code are available by contacting the corresponding author JG by email.

3. Results

3.1. Behavioral data

3.1.1. Linear models

Repeated measures ANOVA on the proportion of choice (Fig. 1b) highlighted a main effect of Emotion, [$F_{(1,25)}$ = 13.77, p = .001, η^2_G = 0.10], as well as a main effect of Intensity, [$F_{(1,25)}$ = 25.02, p < .001, η^2_G = 0.17], respectively indicating that the proportion of "away" responses was greater for anger than for fear, and for high vs. low emotional intensity. The interaction between the two factors was also significant, [$F_{(1,25)}$ = 16.06, p < .001, η^2_G = 0.06]. Paired *t*-tests and Cohen's d effect sizes showed that the difference between high vs. low emotional intensity was large for anger trials, [$t_{(25)}$ = 5.36, p < .001, d = 1.09, d_{CI} = 0.57–1.60], and medium for fear ones, [$t_{(25)}$ = 2.49, p = .020, d = 0.54, d_{CI} = 0.07–1.00], suggesting that the effect of Intensity was stronger for anger vs. fear.

Repeated measures ANOVA on RTs (Fig. 1b) revealed a main effect of Intensity, [$F_{(1,25)}$ = 5.38, p = .029, η^2_G = 0.0003], further characterized by an interaction between Intensity and Side, [$F_{(1,25)}$ = 14.37, p < .001, η^2_G = 0.0005]. The difference between high vs. low emotional intensity in RTs was negligible but significant when participants chose to avoid the emotional individual, [$t_{(25)}$ = 3.78, p < .001, d = 0.08, d_{CI} = 0.04–0.12], and not significant when the participants decided to approach, [$t_{(25)}$ = −0.87, p = .395, d = −0.01, d_{CI} = −0.04–0.02], suggesting that the level of emotional intensity impacted RTs only for away and not for toward responses.

No main effects or interactions were significant for movement times (all F's < 1.38, all p's > .251).

The comparison of mixed models for both the proportion of choice and the RTs confirmed the above-mentioned ANOVA results. In particular, for the mixed logistic models on choice, the Emotion model, the Intensity model and the Emotion by Intensity model were more probable than their respective restricted models, respectively of 29.46, 58.65 and 15.37 bits of evidence (see Supplemental Material, Table S1). Similarly, for RTs, the model including the Side by Intensity interaction was favored compared to its restricted one by 5.01 bits of evidence (see Table S2).

3.1.2. Drift diffusion models

The mean AIC indicated that overall the model where only the drift rate (v) varied over the Emotion and Intensity factors (model 3, mean AIC = −450.34) fitted better the data compared to the model where only the pre-decisional bias (starting point z) varied over the Emotion and Intensity factors (model 2, mean AIC = −447.68). Importantly, this model also fitted better the data compared to the model were both starting point and drift rate varied over the Emotion and Intensity factors (model 4, mean AIC = −446.98) as well as compared to the null model (model 1, mean AIC = −447.79). Critically, the exceedance probability of model 3 was of 0.9997 compared to model 2, of 1 compared to model 4 and of 0.6603 compared to model 1. Overall, there is good support for the hypothesis that model 3 better explains the data compared to the other models tested. Finally, the fit of the winning model (model 3) was also assessed visually, ensuring that it could reproduce the main features of the data (see Fig. S1).

Repeated-measures ANOVA on the drift rate parameter extracted from the winning model 3 (Fig. 2) highlighted a significant main effect

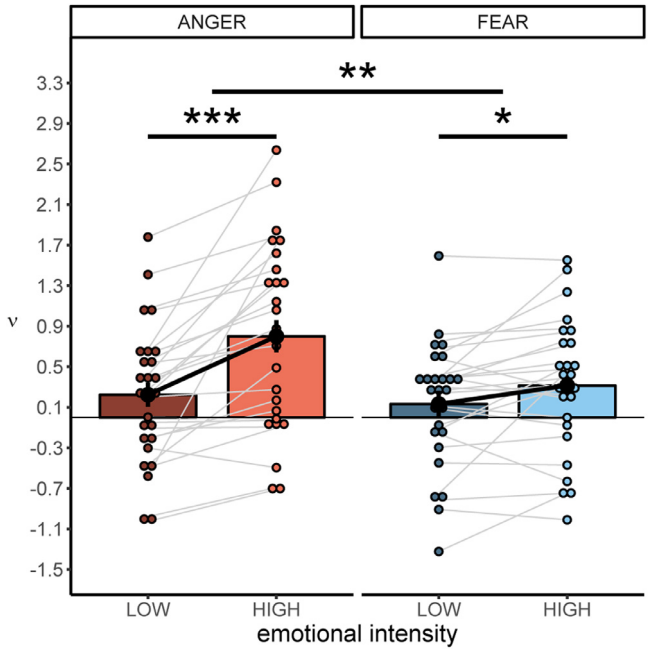


Fig. 2. Drift rate. Black dots and vertical lines represent the mean and the within-subject confidence interval of the mean values for each subject (smaller colored dots). Thinner grey lines connect the values in the "low" and "high" emotional intensity conditions within subjects. *** = $p < .001$; ** = $p < .01$; * = $p < .05$; n.s. = $p > .05$.

of Emotion, [$F_{(1,25)} = 17.69$, $p = .001$, $\eta^2_G = 0.04$], indicating that evidence accumulation for away vs. toward responses was higher for anger than for fear trials. The main effect of Intensity [$F_{(1,25)} = 21.53$, $p < .001$, $\eta^2_G = 0.06$] indicated that evidence accumulation for away vs. toward responses increased with emotional intensity. A significant interaction between Emotion and Intensity also emerged, [$F_{(1,25)} = 12.45$, $p = .002$, $\eta^2_G = 0.02$]: paired *t*-tests and Cohen's d effect sizes showed that the difference between high vs. low emotional intensity was medium for anger trials, [$t_{(25)} = 4.92$, $p < .001$, $d = 0.67$, $d_{CI} = 0.37–0.97$], and small for fear ones, [$t_{(25)} = 2.37$, $p = .03$, $d = 0.28$, $d_{CI} = 0.04–0.52$], suggesting that the effect of Intensity on evidence accumulation was stronger for anger vs. fear.

3.2. EEG data

3.2.1. Anger trials: value encoding

The cluster-based analysis on the betas for the stimulus-locked EEG encoding of value difference for anger trials revealed a significant negative centro-frontal cluster (cluster *t*-value$_{sum}$ = −479.03, $p_{corr} = .033$, time window = 160–248 ms, electrodes = F1, F3, FC3, FC1, C1 C3, CP3, CP1, P1, PO3, Pz, CPz, AFz, Fz, F2, F4, F6, FC4, FC2, FCz, Cz, C2, C4, CP4, CP2, P2, P4; Fig. 3, upper part). The average beta coefficients in the significant cluster for anger differed significantly from those for fear over the same time-window and electrodes ($t_{(25)} = −2.82$, $p = .009$). Source analysis revealed that this effect was mainly associated with activation of bilateral orbitofrontal cortex (OFC) and the left ventromedial prefrontal cortex (vmPFC). Fig. S2 illustrates the same results in a different way: Anger trials that showed strong EEG negativity (i.e., high encoding) on average in the significant spatio-temporal cluster (within-subject median split) were associated with a higher probability to result in an avoidance response, and this more strongly with increasing emotional intensity. Importantly, such early stimulus-locked neural encoding of value difference peaked before selective attention was allocated to emotional displays, as indicated by the peak of the Early Posterior Negativity of the Event-Related Potentials around 290 ms (Figs. S3 and S4).

The cluster-based analysis on the betas for the feedback-locked EEG encoding of value difference for anger trials revealed a significant negative centro-parietal cluster (cluster *t*-value$_{sum}$ = −555.67, $p_{corr} = .026$, time window = 145–265 ms, electrodes = FC1, C1, CP3, CP1, P1, P3, PO7, PO3, Oz, POz, Pz, CPz, Cz, C2, CP4, CP2, P2, P4, P6, PO8, O2; Fig. 3, lower part). The average beta coefficients in the significant cluster for anger differed significantly from those for fear over the same time-window and electrodes ($t_{(25)} = −4.77$, $p < .001$). Source analysis revealed that this effect was mainly associated with activation of bilateral posterior cingulate cortex (PCC) and less pronounced activation was also found in the OFC and the vmPFC.

3.2.2. Fear trials: value encoding

Neither the stimulus-locked (cluster *t*-value$_{sum}$ = 128.49, $p_{corr} = .99$) nor the feedback-locked (cluster *t*-value$_{sum}$ = −237.24, $p_{corr} = .712$) analyses revealed any significant cluster for value difference encoding for fear trials.

3.3. EEG-behavior

Correlation analyses revealed that the average stimulus-locked EEG activity over the value difference encoding significant cluster for anger trials correlated positively with the non-decision time ($\rho = 0.47$, $p = .017$; Fig. 4b). In other words, subjects for which stimulus encoding (and/or response preparation) took longer showed reduced early EEG encoding of value. No correlations emerged with the difference in drift rate between high and low intensity trials ($\rho = 0.08$, $p = .686$).

On the other hand, the feedback-locked EEG cluster negatively correlated with the high minus low intensity difference in drift rate ($\rho = −0.44$, $p = .026$). In other words, higher evidence accumulation for high vs. low intensity anger trials was associated to a stronger EEG encoding of value difference during feedback, after response termination. Feedback-locked activity did not correlate with non-decision time ($\rho = −0.00$, $p = .984$).

4. Discussion

Recent evidence from both animal and human experiments suggests that, in most real-life situations, approach/avoidance tendencies to emotional stimuli can be under the influence of decisional, goal-directed, processes (Moors et al., 2017), as opposed to being solely the result of automatic pre-decisional motor ones. Here, we aimed at providing a precise understanding of the cognitive and neural mechanisms underlying spontaneous approach/avoidance decisions in a realistic context, offering competing targets for action in the presence of a threat-related signal (an angry or fearful individual). Our results are threefold. First, participant's choices to avoid were more frequent in the presence of unambiguously threatening individuals (anger) and of emotional displays of high intensity. Second, participants' choices were accounted for by a process of value-based evidence accumulation, more than by a pre-decisional bias. Third, for unambiguous threat-signaling anger displays, neural encoding of the difference in value between chosen vs. unchosen action options was observed between stimulus presentation and movement initiation. Altogether, the present findings convincingly support the underestimated influence of goal-directed processes on action decisions to threat-signaling expressions.

To tackle the issue of action-related decisions in response to threat-signaling emotional displays, we employed a task in which participants were asked to freely decide between two competing targets for action, in the presence of a neutral individual and another one displaying a threat-related expression (Vilarem et al., 2019). This free-choice paradigm differs in several ways from existing forced-choice compatibility tasks (e.g. AAT). In these tasks, participants are instructed to

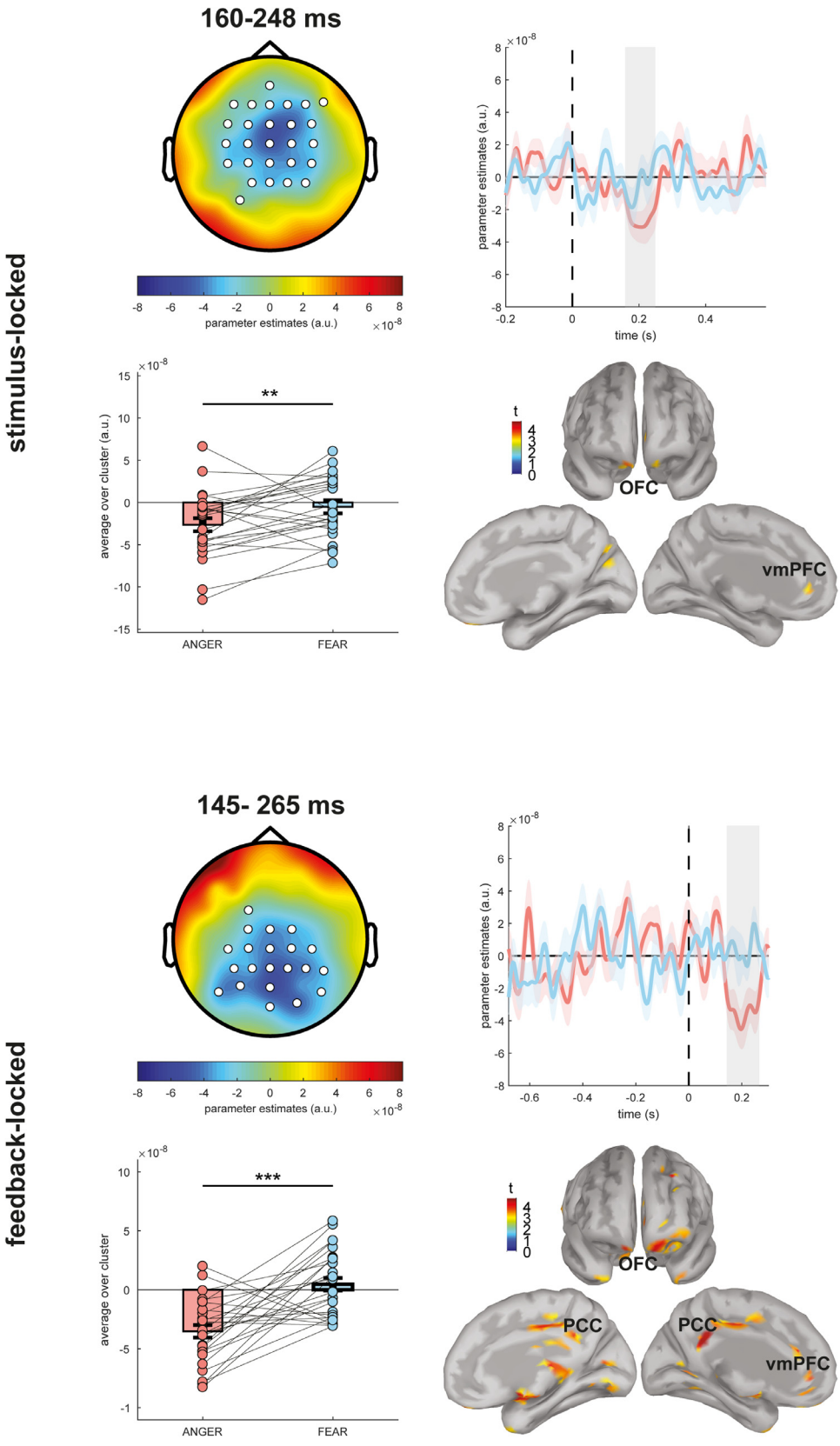


Fig. 3. Value encoding in anger trials. The top part of the figure represents the results for the stimulus-locked analysis, while the bottom part results for the feedback-locked one. For both top and bottom parts, topo-plots (average of the parameter estimates of the GLM over time) and time-courses (average over sensors, mean ± se) of the significant cluster for anger trials (red), with the corresponding (non-significant) parameter estimates for fear trials (blue) are shown. Bar plots (mean ± se) show the average over time and sensors between anger and fear in the same cluster. Brain activations represents the results from a *t*-test against zero on the average over the clusters' significant time-window of the parameter estimates at the source level for anger trials ($p < .005_{unc}$, min voxel = 2). vmPFC = ventromedial prefrontal cortex, OFC = orbitofrontal cortex, PCC = posterior cingulate cortex. *** < .001, ** < .01,* < .05.

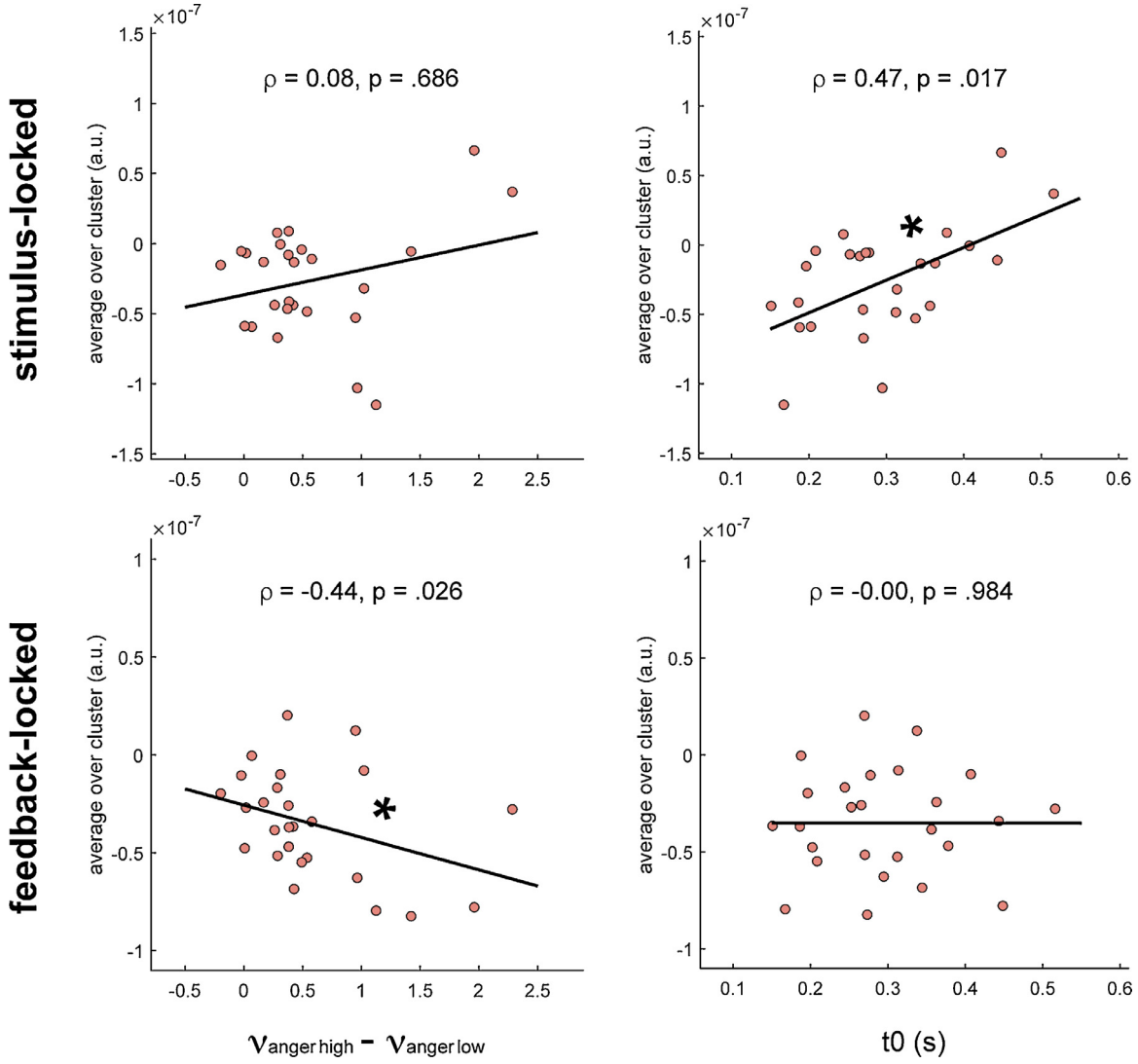


Fig. 4. EEG-DDM correlations for anger trials. On the y axis clusters means are represented for each subject, either in the stimulus-locked (first row) and feedback-locked (second row) clusters encoding for value. The *x* axis are represents the difference in drift rate between high and low emotional intensity trials, left column) and the non-decision time (*t*0; right column).

respond following an explicit rule, which leads them to approach positive stimuli/avoid negative ones in compatible trials and to approach negative stimuli/avoid positive ones in incompatible trials. Faster reaction times are typically observed for compatible vs. incompatible trials (e.g., Chen and Bargh, 1999; De Houwer et al., 2001; Marsh et al., 2005; Solarz, 1960). This effect is often interpreted as the result of a conflict between the instructed action and the action tendency automatically elicited by the emotional stimulus, in the form of a pre-decisional motor bias.

Such interpretation has nonetheless been recently challenged (Eder and Hommel, 2013; Moors et al., 2019, 2017; Rotteveel and Phaf, 2004; Schlund et al., 2016). Firstly, there is evidence that compatibility effects can depend on the ultimate goal of the action being performed: for instance, individuals respond faster in order to ultimately avoiding negative stimuli, even when this implies initially approaching them (Reichardt, 2018). Secondly, a number of other factors were shown to have an influence on how subjects respond to the very same emotional stimulus, such as the explicit label ("approach"/"avoidance") assigned to the movement (Kozlik et al., 2015; Laham et al., 2015), subject's self-representation in space (Seibt et al., 2008) and the presence of other emotional stimuli in the task (Paulus and Wentura, 2016). Each of these factors might explain why AAT paradigms yielded some discrepant results regarding action tendencies to angry and fearful expressions (Bossuyt et al., 2014; Krieglmeyer et al., 2013; Marsh et al., 2005; Paulus and Wentura, 2016; Wilkowski and Meier, 2010). Furthermore, it suggests that approach and avoidance tendencies are not always automatically evoked by the stimulus in a pre-decisional manner but might depend, at least in part, on goal-directed decisional processes (Moors et al., 2017; Moors and Fischer, 2018).

In order to investigate to what extent goal-directed processes contribute to approach-avoidance decisions in emotion, our paradigm had the advantageous characteristics that participants where free to choose among alternatives in a scene representing an everyday environment, the waiting room, without the constraint of instructions, arbitrary movements or response labels. In our opinion, this allowed to simulate more closely how in everyday life different alternatives for action compete to determine spontaneous approach-avoidance responses to emotion displays. Our results indicate that participants exhibited the expected preference to choose the chair that allowed avoiding individuals displaying threat-signaling expressions, in line with previous AAT results. Nonetheless, contrary to what is typically observed in AAT studies, we did not observe a decrease in reaction times when avoiding high vs. low threatening stimuli, a result which appears inconsistent with the existence of a pre-decisional motor bias.

Participant's approach/avoidance decisions were influenced by the presence of facial displays of emotion, as a function of their implied threat, i.e. their behavioral relevance to them (Sander et al., 2007). Replicating previous findings from our team, avoidance was more common when facing angry compared to fearful individuals (Vilarem et al., 2019) and increased as a function of the intensity of the expressed emotion. Although angry and fearful displays are of negative valence, they differ in their social meaning and therefore in their action requests to the perceiver (Horstmann, 2003). Angry expressions are clear signals of impending verbal or physical assault (Sander et al., 2007), which in most contexts leads to avoidance. Fearful expressions, in contrast, signal both the presence of a potential danger (Paulus and Wentura, 2016) and a need for affiliation (Marsh et al., 2005) and are therefore more ambiguous in terms of avoidance and approach decisions. We propose that the above-described influence of threat-related expressions on participant's approach/avoidance decisions is mediated by changing the expected value of each available action option (Vilarem et al., 2019). Crucially, as our task did not use monetary or point incentives (explicit rewards), we anticipated that being seated next to or far from an emotional individual would be a high motivational outcome per se, even in the context of a laboratory task. Therefore, in agreement with a goal-directed perspective, anger, and to a lesser extent fear, would increase the value of the action leading to the most desirable outcome, i.e. threat avoidance.

A goal-directed perspective entails that decisions between action options depend on valuation and comparison between available options to generate a choice (e.g., Wunderlich et al., 2009; Xie and Padoa-Schioppa, 2016), a mechanism characterized as evidence accumulation (e.g., Polanía et al., 2014). As multiple cognitive processes can give rise to similar patterns of participant performance, we fitted drift-diffusion models to participant choice behavior and RTs. We found a higher rate of evidence accumulation, i.e. higher value estimates, when participants spontaneously decided to avoid (compared to approach) individuals displaying angry expressions (compared to fearful ones), especially at high emotional intensity. While replicating previous findings from computational modeling of participants' non-spontaneous approach-avoidance responses and RTs during AAT (Krypotos et al., 2015; Tipples, 2018), we suggest here that a change in evidence accumulation provides strong evidence for a rapid and implicit decisional process, underlying approach/avoidance responses to emotional stimuli, rather than a pre-decisional one.

The early neural encoding of the key decision variable guiding choice, i.e. the difference in value between a choice taken and a choice untaken (Papageorgiou et al., 2017) when facing unambiguous angry displays, further confirms our interpretation. This early value difference signal was observed in fronto-central medial electrodes, around 200 ms after the onset of a scene with an angry individual and before movement initiation. Importantly, this implies that EEG activity in this spatiotemporal cluster predicted subsequent participant choices on a trial-by trial basis. Furthermore, the encoding of the value difference was sourced in the ventromedial prefrontal and orbitofrontal cortices (vmPFC/OFC). Both regions have been identified as especially important in value-based decisions, notably contributing to encode value differences between alternatives (e.g., Boorman et al., 2009; FitzGerald et al., 2009; Lim et al., 2011; Rushworth et al., 2011; Hunt et al., 2012, 2013, 2015; Levy and Glimcher, 2012; Jocham et al., 2014; Setogawa et al., 2019), a signal found to be predictive of subsequent choices (e.g., Howard and Kahnt, 2017). Here, in accordance with previous findings, the stronger the neural signal in this early spatiotemporal cluster, the higher the probability that the participant chooses the action leading to the most desirable outcome, i.e. avoiding individuals displaying angry expressions. Interestingly, these results are consistent with a previous AAT study, investigating the neural correlates of approach vs. avoidance in alcohol-dependent patients compared to controls (Wiers et al., 2014). In this study, BOLD activations in the vmPFC and nucleus accumbens were stronger when approaching vs. avoiding each group's most desirable outcome, i.e., alcohol for patients and soft drinks for controls. Our results replicate and extend these findings, endorsing the role of the value comparison process in driving approach/avoidance decisions in the presence of threatening individuals.

The early neural encoding of value difference peaked before selective attention was allocated to emotional displays. This is in agreement with the recent proposal that flexible decisional processes can include implicit forecasting of action outcomes (LeDoux and Daw, 2018), thus combining speed with optimality, contrary to the common intuition that equates goal-directed behaviors to slow and costly responses (Hommel and Wiers, 2017; Moors, 2017; Moors et al., 2019; Moors and Fischer, 2018). For instance, it has been recently shown that humans are able to rapidly respond, within around 200 ms, to evolving sensory information in a manner consistent with value-based decision-making (Carroll et al., 2019). In our study, we confirm and extend these results, by showing that emotional displays in the environment play an important role in this rapid value-based arbitration between action alternatives. Of note, the inverse correlation between the non-decision time parameter, which captures stimulus encoding and response execution, and the early value difference signal suggests that the longer participants took to process the scene and prepare their responses, the less efficient their value comparison process. This implies that, despite being extremely rapid, implicit decisional processes still require some availability of cognitive and time resources (Marien et al., 2012).

Finally, in agreement with our assumption that being seated far from an angry individual was a highly desirable outcome, the value difference between chosen and unchosen options modulated the EEG signal around 200 ms after choice feedback (i.e., a picture of the participant on the chosen chair). Furthermore, the more efficient the participant's decision process was in the presence of angry expressions of high compared to low intensity (i.e. higher rate of evidence accumulation), the stronger their neural encoding of value difference during feedback. This effect was sourced in the posterior cingulate cortex and in vmPFC/OFC. Activity in these brain areas which, together with the ventral striatum, constitute the brain's valuation system, not only "scales with the subjective value of the available alternatives during choice" (Bartra et al., 2013, p. 412), but "also responds when reward is received, implicating a common set of regions in the evaluation of both prospects and outcomes" (Bartra et al., 2013, p. 412) (see also Mccoy et al., 2003; Strait et al., 2014). Our correlational findings are further consistent with the observation that the more the predicted consequences of a choice matches its real outcome, the more the vmPFC is active (Blanchard and Gershman, 2018).

The conclusions of this study ought to be interpreted in light of some limitations. Theoretically, while our results support the idea that the relative value of each action was indeed computed before choice, they do not guarantee that at each trial subjects were responding based on the forecasted consequence of their action. In the literature, truly goal-directed actions are supposed to be sensitive to changes in (i) the causal relationship to their consequences and (ii) the value of those consequences (Balleine and Dickinson, 1998). Common tests to assess whether an action demonstrates sensitivity to these changes involve devaluation procedures, which modify the value of the associated outcome (e.g. satiation for food rewards) and contingency degradation procedures that modify the contingency between action and outcome (e.g., lowering the probability that by pressing a lever, one will obtain food).

In future versions of the present task, devaluation could be implemented by building trials in which, after the choice, the emotional expressions of the seated individuals change in the feedback phase, thus rendering the consequence of the action less predictable. If subjects' avoidance really depends on forecasted consequence of each action possibility, it should strongly diminish in a context of low predictability. On the other hand, contingency degradation could be implemented by unpredictably switching mouse coordinates on the *x*-axis in some trials, to test whether, in the condition of total unpredictability, subjects' initial hand movement would still be most of the times in the direction of

threat avoidance. We believe that this is an interesting research line for future studies.

From a methodological point of view, the very nature of the feedback employed in the present experiment ought to be studied in more details. On the basis of previous results showing that approach-avoidance of emotional stimuli can be motivated by the anticipated desirable consequences associated with the actions themselves (Eder et al., 2015), we presented participants' own picture on the chosen chair after a proper movement, to embody the consequence of participant's choice and strengthen the impression of an accomplished movement. Our EEG results indeed show that participants took into account this feedback, as they encoded the value associated with the chosen vs. unchosen option during the feedback phase, similarly to what happens in classic preference-based choice (Bartra et al., 2013). Nonetheless, we never directly tested whether the behavioral results would change in the absence of the feedback, leaving open the possibility that this explicit feedback was not necessary for participants to forecast the consequences of their actions. Moreover, while EEG allowed to precisely test our hypotheses on the temporal characteristics of brain activation during choice, our conclusions about brain sources are inherently limited in terms of spatial localization and ought to be substantiated by future studies using more spatially accurate imaging techniques. Finally, this study largely replicated results regarding the proportion of choice found in our previous work using the same paradigm (Vilarem et al., 2019). Results on RT seem overall smaller and more variable, therefore they ought to be interpreted with caution, awaiting for future replications with larger sample sizes.

5. Conclusions

Overall, the present study strongly supports the idea that approach-avoidance tendencies to emotion depend, at least in part, on implicit value-based decisions. Such a conclusion, if substantiated, might have important theoretical and clinical implications. Theoretically, the fact that an emotion-based process of value attribution can influence action selection very rapidly, and possibly outside consciousness (e.g., Pessiglione et al., 2007; Wimmer and Shohamy, 2012), reinforces the idea that our implicit motivations, goals and expectancies about our interactions with others are likely to have a profound impact on how we spontaneously navigate our socio-emotional environment. This emotion-based process of value attribution seems far from being highly demanding cognitively and computationally, and the way social expectancies are built likely depends on several types of learning and memory processes (Amodio, 2019). Explaining how precisely these different processes contribute to implicit value assignment is an exciting topic for future research.

We agree with the idea that the relationship between emotion and action goes beyond simple stimulus-driven pre-decisional reactions, such as species typical reactions and habits (LeDoux and Daw, 2018), and the way we inhibit or refine them (Moors et al., 2017). This broadened perspective might impact how we deal with emotional disturbance in psychopathology. Indeed, it would speak in favor of therapeutic approaches aimed at understanding and eventually modifying implicit expectancies of future action outcomes in response to emotional signals, rather than focusing on instantiating new stimulus-driven associations or promoting explicit top-down control.

6. Funding

This work was supported by FRM Team DEQ20160334878; and Fondation de France 00100076; and INSERM; and ENS; and the French National Research Agency under Grants ANR-10-LABX-0087 IEC and ANR-17-EURE-0017 FrontCog. R.M. was supported by Fondation de France and FRM postdoctoral fellowships and benefited from a Gretty Mirdal Junior Fellowship at the Paris Institute for Advanced Study (France). The funding sources had no role in influencing any research stage.

CRediT author statement

Rocco Mennella: Methodology, Software, Formal analysis, Writing - Original Draft. **Emma Vilarem:** Conceptualization, Methodology, Software, Investigation, Writing - Review & Editing. **Julie Grèzes:** Conceptualization, Supervision, Project administration, Resources, Funding acquisition, Writing - Review & Editing.

Declaration of Competing Interest

None.

Acknowledgments

The authors thank Michèle Chadwick for carefully proofreading this manuscript, Tarryn Balsdon for advices regarding DDMs and Damiano Azzalini for his useful comments.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2020.117199.

References

Amodio, D.M., 2019. Social cognition 2.0: an interactive memory systems account. Trends Cogn. Sci. 23, 21–33. doi:10.1016/j.tics.2018.10.002.

Bach, D.R., Dayan, P., 2017. Algorithms for survival: A comparative perspective on emotions. Nat. Rev. Neurosci. 18, 311–319. doi:10.1038/nrn.2017.35.

Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37, 407–419. doi:10.1016/S0028-3908(98)00033-1.

Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76, 412–427. doi:10.1016/j.neuroimage.2013.02.063.

Blanchard, T.C., Gershman, S.J., 2018. Pure correlates of exploration and exploitation in the human brain. Cogn. Affect. Behav. Neurosci. 18, 117–126. doi:10.3758/s13415-017-0556-2.

Boorman, E.D., Behrens, T.E.J., Woolrich, M.W., Rushworth, M.F.S., 2009. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron 62, 733–743. doi:10.1016/j.neuron.2009.05.014.

Bossuyt, E., Moors, A., De Houwer, J., 2014. On angry approach and fearful avoidance: The goal-dependent nature of emotional approach and avoidance tendencies. J. Exp. Soc. Psychol. doi:10.1016/j.jesp.2013.09.009.

Bramson, B., Jensen, O., Toni, I., Roelofs, K., 2018. Cortical oscillatory mechanisms supporting the control of human social–emotional actions. J. Neurosci. 38, 5739–5749. doi:10.1523/JNEUROSCI.3382-17.2018.

Cain, C.K., 2019. Avoidance problems reconsidered. Curr. Opin. Behav. Sci. 26, 9–17. doi:10.1016/j.cobeha.2018.09.002.

Carroll, T.J., McNamee, D., Ingram, J.N., Wolpert, D.M., 2019. Rapid visuomotor responses reflect value-based decisions. J. Neurosci. 39, 3906–3920. doi:10.1523/JNEUROSCI.1934-18.2019.

Chen, M., Bargh, J.A., 1999. Consequences of automatic evaluation: immediate behavioral predispositions to approach or avoid the stimulus. Personal. Soc. Psychol. Bull. 25, 215–224. doi:10.1177/0146167299025002007.

De Houwer, J., Crombez, G., Baeyens, F., Hermans, D., 2001. On the generality of the affective Simon effect. Cogn. Emot. 15, 189–206. doi:10.1080/02699930125883.

Delorme, A., Makeig, S., 2004. EEGLAB: an open sorce toolbox for analysis of single-trail EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21. doi:10.1016/j.jneumeth.2003.10.009.

Dezecache, G., Jacob, P., Grèzes, J., 2015. Emotional contagion: Its scope and limits. Trends Cogn. Sci. doi:10.1016/j.tics.2015.03.011.

Eder, A.B., Hommel, B., 2013. Anticipatory control of approach and avoidance: An ideomotor approach. Emot. Rev. 5, 275–279. doi:10.1177/1754073913477505.

Eder, A.B., Rothermund, K., De Houwer, J., Hommel, B., 2015. Directive and incentive functions of affective action consequences: an ideomotor approach. Psychol. Res. 79, 630–649. doi:10.1007/s00426-014-0590-4.

El Zein, M., Wyart, V., Grèzes, J., 2015. Anxiety dissociates the adaptive functions of sensory and motor response enhancements to social threats. Elife 4, 1–22. doi:10.7554/eLife.10274.001.

Evans, D.A., Stempel, A.V., Vale, R., Branco, T., 2019. Cognitive control of escape behaviour. Trends Cogn. Sci. 23, 334–348. doi:10.1016/j.tics.2019.01.012.

Fang, Y., 2011. Asymptotic equivalence between cross-validations and akaike information criteria in mixed-effects models. Data Sci. 9, 15–25.

FitzGerald, T.H.B., Seymour, B., Dolan, R.J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. 29, 8388–8395. doi:10.1523/JNEUROSCI.0717-09.2009.

Frijda, N.H., 1986. The Emotions. Cambridge University Press, Cambridge.

Glover, S., Dixon, P., 2004. Likelihood ratios: A simple and flexible statistic for empirical psychologists. Psychon. Bull. Rev. 11, 791–806. doi:10.3758/BF03196706.
Grèzes, J., Valabrègue, R., Gholipour, B., Chevallier, C., 2014. A direct amygdala-motor pathway for emotional displays to influence action: A diffusion tensor imaging study. Hum. Brain Mapp. 35, 5974–5983. doi:10.1002/hbm.22598.
Hammer, J.L., Marsh, A.A., 2015. Why do fearful facial expressions elicit behavioral approach? Evidence from a combined approach-avoidance implicit association test. Emotion 15, 223–231. doi:10.1037/emo0000054.
Hashemi, M.M., Gladwin, T.E., de Valk, N.M., Zhang, W., Kaldewaij, R., van Ast, V., Koch, S.B.J., Klumpers, F., Roelofs, K., 2019. Neural dynamics of shooting decisions and the switch from freeze to fight. Sci. Rep. 9, 1–10. doi:10.1038/s41598-019-40917-8.
Hommel, B., Wiers, R.W., 2017. Towards a unitary approach to human action control. Trends Cogn. Sci. 21, 940–949. doi:10.1016/j.tics.2017.09.009.
Horstmann, G., 2003. What do facial expressions convey: Feeling states, behavioral intentions, or action requests? Emotion 3, 150–166. doi:10.1037/1528-3542.3.2.150.
Howard, J.D., Kahnt, T., 2017. Identity-specific reward representations in orbitofrontal cortex are modulated by selective devaluation. J. Neurosci. 37, 2627–2638. doi:10.1523/JNEUROSCI.3473-16.2017.
Hunt, L.T., Behrens, T.E., Hosokawa, T., Wallis, J.D., Kennerley, S.W., 2015. Capturing the temporal evolution of choice across prefrontal cortex. Elife 4, 1–25. doi:10.7554/elife.11945.
Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F.S., Behrens, T.E.J., 2012. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 15, 470–476. doi:10.1038/nn.3017.
Hunt, L.T., Woolrich, M.W., Rushworth, M.F.S., Behrens, T.E.J., 2013. Trial-type dependent frames of reference for value comparison. PLoS Comput. Biol. 9. doi:10.1371/journal.pcbi.1003225.
Jocham, G., Hunt, L.T., Near, J., Behrens, T.E., 2012. A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. Nat. neurosci 15, 960–961. doi:10.1038/nn.3140.
Jones, C.R., Vilensky, M.R., Vasey, M.W., Fazio, R.H., 2013. Approach behavior can mitigate predominately univalent negative attitudes: Evidence regarding insects and spiders. Emotion 13, 989–996. doi:10.1037/a0033164.
Keltner, D., Haidt, J., 1999. Social functions of emotions at four levels of analysis. Cogn. Emot. 13, 505–521. doi:10.1080/026999399379168.
Kozlik, J., Neumann, R., Lozo, L., 2015. Contrasting motivational orientation and evaluative coding accounts: On the need to differentiate the effectors of approach/avoidance responses. Front. Psychol. 6, 1–10. doi:10.3389/fpsyg.2015.00563.
Krieglmeyer, R., De Houwer, J., Deutsch, R., 2013. On the nature of automatically triggered approach-avoidance behavior. Emot. Rev. 5, 280–284. doi:10.1177/1754073913477501.
Krypotos, A.M., Beckers, T., Kindt, M., Wagenmakers, E.J., 2015. A Bayesian hierarchical diffusion model decomposition of performance in approach – Avoidance tasks. Cogn. Emot. 29, 1424–1444. doi:10.1080/02699931.2014.985635.
Laham, S.M., Kashima, Y., Dix, J., Wheeler, M., 2015. A meta-analysis of the facilitation of arm flexion and extension movements as a function of stimulus valence. Cogn. Emot. 29, 1069–1090. doi:10.1080/02699931.2014.968096.
Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1990. Emotion, attention, and the startle reflex. Psychol. Rev. 97, 377–395. doi:10.1037/0033-295X.97.3.377.
Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., van Knippenberg, A., 2010. Presentation and validation of the radboud faces database. Cogn. Emot. doi:10.1080/02699930903485076.
Lawrence, M., 2016. Package "ez." R Top. Doc.
LeDoux, J.E., 2012. Rethinking the emotional brain. Neuron 73, 653–676. doi:10.1016/j.neuron.2012.02.004.
LeDoux, J.E., 1996. The Emotional Brain: The Mysterious Underpinnings of Emotional Life. Simon & Schuster.
LeDoux, J.E., Daw, N.D., 2018. Surviving threats: Neural circuit and computational implications of a new taxonomy of defensive behaviour. Nat. Rev. Neurosci. doi:10.1038/nrn.2018.22.
Lerche, V., Voss, A., 2016. Model complexity in diffusion modeling: benefits of making the model more parsimonious. Front. Psychol. 7. doi:10.3389/fpsyg.2016.01324.
Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. Curr. Opin. Neurobiol. 22, 1027–1038. doi:10.1016/j.conb.2012.06.001.
Lim, S.L., O'Doherty, J.P., Rangel, A., 2011. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. J. Neurosci. 31, 13214–13223. doi:10.1523/JNEUROSCI.1246-11.2011.
Marien, H., Custers, R., Hassin, R.R., Aarts, H., 2012. Unconscious goal activation and the hijacking of the executive function. J. Pers. Soc. Psychol. 103, 399–415. doi:10.1037/a0028955.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190. doi:10.1016/j.jneumeth.2007.03.024.
Marsh, A.A., Ambady, N., Kleck, R.E., 2005. The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. Emotion 5, 119–124. doi:10.1037/1528-3542.5.1.119.
McCoy, A.N., Crowley, J.C., Haghighian, G., Dean, H.L., Platt, M.L., 2003. Saccade reward signals in posterior cingulate cortex. Neuron 40, 1031–1040. doi:10.1016/S0896-6273(03)00719-0.
Moors, A., 2017. The integrated theory of emotional behavior follows a radically goal-directed approach. Psychol. Inq. 28, 68–75. doi:10.1080/1047840x.2017.1275207.
Moors, A., Boddez, Y., De Houwer, J., 2017. The power of goal-directed processes in the causation of emotional and other actions. Emot. Rev. 9, 310–318. doi:10.1177/1754073916669595.
Moors, A., Fini, C., Everaert, T., Bardi, L., Bossuyt, E., Kuppens, P., Brass, M., 2019. The role of stimulus-driven versus goal-directed processes in fight and flight tendencies measured with motor evoked potentials induced by Transcranial Magnetic Stimulation. PLoS One 14, e0217266. doi:10.1371/journal.pone.0217266.
Moors, A., Fischer, M., 2018. Demystifying the role of emotion in behaviour: toward a goal-directed account. Cogn. Emot. 0, 1–7. doi:10.1080/02699931.2018.1510381.
Öhman, A., 1986. Face the beast and fear the face: animal and social fears as prototypes for evolutionary analyses of emotion. Psychophysiology 23, 123–145. doi:10.1111/j.1469-8986.1986.tb00608.x.
Papageorgiou, G.K., Sallet, J., Wittmann, M.K., Chau, B.K.H., Schüffelgen, U., Buckley, M.J., Rushworth, M.F.S., 2017. Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. Nat. Commun. 8, 1886. doi:10.1038/s41467-017-01833-5.
Paré, D., Quirk, G.J., 2017. When scientific paradigms lead to tunnel vision: lessons from the study of fear. NPJ Sci. Learn. 2, 6. doi:10.1038/s41539-017-0007-4.
Patron, E., Mennella, R., Messerotti Benvenuti, S., Thayer, J.F., 2019. The frontal cortex is a heart-brake: Reduction in delta oscillations is associated with heart rate deceleration. NeuroImage 188, 403–410. doi:10.1016/j.neuroimage.2018.12.035.
Paulus, A., Wentura, D., 2016. It depends: Approach and avoidance reactions to emotional expressions are influenced by the contrast emotions presented in the task. J. Exp. Psychol. Hum. Percept. Perform. 42, 197–212. doi:10.1037/xhp0000130.
Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., Frith, C.D., 2007. How the brain translates money into force: a neuroimaging study of subliminal motivation. Science 316, 904–906. doi:10.1126/science.1140459.
Polanía, R., Krajbich, I., Grueschow, M., Ruff, C.C., 2014. Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. Neuron 82, 709–720. doi:10.1016/j.neuron.2014.03.014.
Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922. doi:10.1162/neco.2008.12-06-420.
Reichardt, R., 2018. Taking a Detour: Affective stimuli facilitate ultimately (Not Immediately) compatible approach-avoidance tendencies. Front. Psychol. 9, 1–8. doi:10.3389/fpsyg.2018.00488.
Rigoux, L., Stephan, K.E., Friston, K.J., Daunizeau, J., 2014. Bayesian model selection for group studies—Revisited. Neuroimage 84, 971–985. doi:10.1016/j.neuroimage.2013.08.065.
Roberts, I.D., Hutcherson, C.A., 2019. Affect and decision making: insights and predictions from computational models. Trends Cogn. Sci. 23, 602–614. doi:10.1016/j.tics.2019.04.005.
Roelofs, K., Minelli, A., Mars, R.B., van Peer, J., Toni, I., 2009. On the neural control of social emotional behavior. Soc. Cogn. Affect. Neurosci. 4, 50–58. doi:10.1093/scan/nsn036.
Rotteveel, M., Phaf, R.H., 2004. Automatic affective evaluation does not automatically predispose for arm flexion and extension. Emotion 4, 156–172. doi:10.1037/1528-3542.4.2.156.
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, 1054–1069. doi:10.1016/j.neuron.2011.05.014.
Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., Scherer, K.R., 2007. Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. Eur. J. Cogn. Psychol. 19, 470–480. doi:10.1080/09541440600757426.
Schlund, M.W., Brewer, A.T., Magee, S.K., Richman, D.M., Solomon, S., Ludlum, M., Dymond, S., 2016. The tipping point: Value differences and parallel dorsal–ventral frontal circuits gating human approach–avoidance behavior. NeuroImage 136, 94–105. doi:10.1016/j.neuroimage.2016.04.070.
Seibt, B., Neumann, R., Nussinson, R., Strack, F., 2008. Movement direction or change in distance? Self- and object-related approach-avoidance motions. J. Exp. Soc. Psychol. 44, 713–720. doi:10.1016/j.jesp.2007.04.013.
Setogawa, T., Mizuhiki, T., Matsumoto, N., Akizawa, F., Kuboki, R., Richmond, B.J., Shidara, M., 2019. Neurons in the monkey orbitofrontal cortex mediate reward value computation and decision-making. Commun. Biol. 2, 1–9. doi:10.1038/s42003-019-0363-0.
Solarz, A.K., 1960. Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. J. Exp. Psychol. 59, 239–245. doi:10.1037/h0047274.
Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. NeuroImage 46, 1004–1017. doi:10.1016/j.neuroimage.2009.03.025.
Strait, C.E., Blanchard, T.C., Hayden, B.Y., 2014. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron 82, 1357–1366. doi:10.1016/j.neuron.2014.04.032.
Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 1–13. doi:10.1155/2011/879716.
Taylor, C.T., Amir, N., 2012. Modifying automatic approach action tendencies in individuals with elevated social anxiety symptoms. Behav. Res. Ther. 50, 529–536. doi:10.1016/j.brat.2012.05.004.
Tipples, J., 2018. Recognising and reacting to angry and happy facial expressions : a diffusion model analysis. Psychol. Res. 83, 37–47 https://doi.org/10.1007/s00426-018-1092-6.
Toschi, N., Duggento, A., Passamonti, L., 2017. Functional connectivity in amygdalar-sensory/(pre)motor networks at rest: new evidence from the Human Connectome Project. Eur. J. Neurosci. 45, 1224–1229. doi:10.1111/ejn.13544.
Vilarem, E., Armony, J.L., Grèzes, J., 2019. Action opportunities modulate attention allocation under social threat. Emotion doi:10.1037/emo0000598.
Voss, A., Voss, J., 2008. A fast numerical algorithm for the estimation of diffusion model parameters. J. Math. Psychol. 52, 1–9. doi:10.1016/j.jmp.2007.09.005.

Voss, A., Voss, J., 2007. Fast-dm: A free program for efficient diffusion model analysis. Behav. Res. Methods 39, 767–775. doi:10.3758/BF03192967.

Voss, A., Voss, J., Lerche, V., 2015. Assessing cognitive processes with diffusion model analyses: A tutorial based on fast-dm-30. Front. Psychol. 6, 1–14. doi:10.3389/fpsyg.2015.00336.

Waller, B.M., Whitehouse, J., Micheletta, J., 2017. Rethinking primate facial expression: A predictive framework. Neurosci. Biobehav. Rev. 82, 13–21. doi:10.1016/j.neubiorev.2016.09.005.

Wiers, C.E., Stelzel, C., Park, S.Q., Gawron, C.K., Ludwig, V.U., Gutwinski, S., Heinz, A., Lindenmeyer, J., Wiers, R.W., Walter, H., Bermpohl, F., 2014. Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh is weak for spirits. Neuropsychopharmacology 39, 688–697. doi:10.1038/npp.2013.252.

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. Psychol. Sci. 22, 490–497. doi:10.1177/0956797611400615.

Wilkowski, B.M., Meier, B.P., 2010. Bring it on: Angry facial expressions potentiate approach-motivated motor behavior. J. Pers. Soc. Psychol 98, 201–210. doi:10.1037/a0017992.

Wimmer, G.E., Shohamy, D., 2012. Preference by association: how memory mechanisms in the hippocampus bias decisions. Science 338, 270–273. doi:10.1126/science.1223252.

Wunderlich, K., Range, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. Proc. Natl. Acad. Sci. U.S.A. 106, 17199–17204. doi:10.1073/pnas.0901077106.

Wyart, V., de Gardelle, V., Scholl, J., Summerfield, C., 2012. Rhythmic fluctuations in evidence accumulation during decision making in the human brain. Neuron 76, 847–858. doi:10.1016/j.neuron.2012.09.015.

Wyart, V., Myers, N.E., Summerfield, C., 2015. Neural mechanisms of human perceptual choice under focused and divided attention. J. Neurosci. 35, 3485–3498. doi:10.1523/JNEUROSCI.3276-14.2015.

Xie, J., Padoa-Schioppa, C., 2016. Neuronal remapping and circuit persistence in economic decisions. Nat. Neurosci. 19, 855–861. doi:10.1038/nn.4300.