OXFORD

Social Cognitive and Affective Neuroscience, 2022, **17**, 683–693

DOI: https://doi.org/10.1093/scan/nsab125
Advance Access Publication Date: 27 November 2021
Original Manuscript

Dynamic computation of value signals via a common neural network in multi-attribute decision-making

Amadeus Magrabi,[1,2] Vera U. Ludwig[1,2,3,4] Christian M. Stoppel,[2] Lena M. Paschke,[1,2,5] David Wisniewski,[6] Hauke R. Heekeren,[1,7] and Henrik Walter[1,2,8]

[1]Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin 10117, Germany
[2]Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin Berlin, Berlin 10117, Germany
[3]Wharton Neuroscience Initiative, University of Pennsylvania, Philadelphia, PA 19104, USA
[4]Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[5]Department of Psychology, Humboldt-Universität zu Berlin, Berlin 12489, Germany
[6]Department of Experimental Psychology, Ghent University, Gent 9000, Belgium
[7]Department of Education and Psychology, Freie Universität Berlin, Berlin 14195, Germany
[8]Berlin Center for Advanced Neuroimaging, Charité—Universitätsmedizin Berlin, Berlin 10119, Germany
Correspondence should be addressed to Amadeus Magrabi, Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin Berlin, Berlin 10117, Germany. E-mail: amadeus.magrabi@gmail.com.

Abstract

Studies in decision neuroscience have identified robust neural representations for the value of choice options. However, overall values often depend on multiple attributes, and it is not well understood how the brain evaluates different attributes and integrates them to combined values. In particular, it is not clear whether attribute values are computed in distinct attribute-specific regions or within the general valuation network known to process overall values. Here, we used a functional magnetic resonance imaging choice task in which abstract stimuli had to be evaluated based on variations of the attributes color and motion. The behavioral data showed that participants responded faster when overall values were high and attribute value differences were low. On the neural level, we did not find that attribute values were systematically represented in areas V4 and V5, even though these regions are associated with attribute-specific processing of color and motion, respectively. Instead, attribute values were associated with activity in the posterior cingulate cortex, ventral striatum and posterior inferior temporal gyrus. Furthermore, overall values were represented in dorsolateral and ventromedial prefrontal cortex, and attribute value differences in dorsomedial prefrontal cortex, which suggests that these regions play a key role for the neural integration of attribute values.

Key words: decision-making; value; attribute; salience; fMRI

Introduction

Valuation is a crucial part of decision-making. To make beneficial choices, available options need to be accurately evaluated, and the ones with the highest value need to be selected. Studies in decision neuroscience have investigated valuation processes extensively and found neural representations of value in the ventromedial prefrontal cortex (vmPFC), posterior cingulate cortex (PCC) and ventral striatum (Bartra *et al.*, 2013; Clithero and Rangel, 2014). Most studies addressed the question as to where overall values of choice options are processed in the brain (Rangel *et al.*, 2008; Kable and Glimcher, 2009). However, overall values are often based on values of different attributes. For example, the overall value of a car can depend on the evaluation of its size, speed or color. In these cases, values of relevant attributes have to be computed separately, before they can be integrated to a combined value that ultimately determines choices. So far, the majority of studies investigated neural representations of overall values (Sanfey *et al.*, 2006; Kable and Glimcher, 2007; Rangel *et al.*, 2008; Levy and Glimcher, 2012), but little is known about the computation and integration of attribute values (Basten *et al.*, 2010; Kahnt *et al.*, 2011; Lim *et al.*, 2013; Suzuki *et al.*, 2017; Vaidya *et al.*, 2018; Pelletier *et al.*, 2021).

With regard to known functional specializations of different brain regions, two hypotheses concerning neuronal attribute valuation are conceivable. On the one hand, attribute values could be computed in distinct, attribute-specific brain regions (Basten *et al.*, 2010; Philiastides *et al.*, 2010; Lim *et al.*, 2013). From this perspective, attribute values are processed within regions that are also specialized in processing objective properties of the particular attributes. For instance, it would be predicted that the fusiform face area, which is known to selectively process faces (Kanwisher *et al.*, 1997), is also responsible for the evaluation of faces. As a

Received: 2 March 2021; **Revised:** 12 October 2021; **Accepted:** 25 November 2021

© The Author(s) 2021. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

result, attribute values for a choice option would be computed in distinct neural regions that highly depend on the particular attribute, which is consistent with studies that found evidence of value correlations in sensory regions (Gold and Shadlen, 2007; Serences, 2008; Persichetti *et al.*, 2015; Hanks and Summerfield, 2017). On the other hand, all attribute-specific value computations could instead be performed within the general valuation network that is known to process overall values (Bartra *et al.*, 2013; Clithero and Rangel, 2014). As such, different attribute values as well as overall values would be processed in a homogeneous and centralized manner via vmPFC, PCC and ventral striatum (Peters and Büchel, 2010; Ludwig *et al.*, 2014).

Here, we used model-based functional magnetic resonance imaging (fMRI) to distinguish between these hypotheses and investigate how attribute values are computed in the brain. Participants were presented with a dot stimulus varying in two constituent perceptual attributes: motion direction and dot color. Each attribute level (i.e. each particular motion direction and each color) was associated with a specific monetary gain or loss. Based on these individual attribute values, participants had to determine the overall value of the stimulus and decide to accept or reject the offer. In addition, we also conducted separate localizer tasks to identify regions specialized in the processing of the physical properties of the attributes, independent of attribute valuation.

Compared to previous studies on attribute valuation (Hare *et al.*, 2009; Basten *et al.*, 2010; Philiastides *et al.*, 2010; Kahnt *et al.*, 2011; Park *et al.*, 2011; Lim *et al.*, 2013; Hutcherson *et al.*, 2015; Suzuki *et al.*, 2017; de Berker *et al.*, 2019; Pelletier *et al.*, 2021), this approach combines two methodological advantages:

(i) Our decision task is based on two well-investigated stimulus attributes, which are processed in separate, well-defined cortical modules, namely Area V5 for motion (Watson *et al.*, 1993) and Area V4 for color (McKeefry and Zeki, 1997). By using independent localizer tasks, we were thus able to specifically address whether valuation of individual attributes proceeds separately within these well-defined perceptual areas, or instead within the network known to compute overall values (comprising vmPFC, PCC and ventral striatum).

(ii) The values assigned to each attribute in our task span a range of both positive and negative values, which allows us to disentangle the effects of value and salience. Salience, in contrast to value, refers to the subjective importance of a stimulus, which ultimately guides the amount of attentional resources deployed to a certain stimulus or event (Maunsell, 2004; Zink *et al.*, 2006; Kahnt and Tobler, 2013). When valuation processes are studied only by means of positive values, salience and value are indistinguishable, as more positive stimuli are also more salient (Litt *et al.*, 2011; Leathers and Olson, 2012; Kahnt *et al.*, 2014; Zhang *et al.*, 2017). For that reason, salience is a common experimental confound in the majority of decision neuroscience studies (O'Doherty, 2014). However, if both positive and negative values are involved, salience- and valuation-related mechanisms of decision processes can easily be distinguished, because stimuli with a high negative value are of low value, but of high salience (as there is a high incentive to avoid them). Hence, our design allows us to dissociate regions that compute the value and salience of each attribute, to provide a more elaborate account of attribute valuation in the human brain.

Methods

Participants

Twenty-five right-handed subjects (14 female; mean age 28.1 ± 4 s.d.) participated in the study. All subjects had normal or corrected-to-normal vision, had no history of psychiatric or neurological illnesses, were free from medication interfering with fMRI performance, were native German speakers, and provided informed consent before participation. All experimental procedures were approved by the local ethics committee. Participants were compensated with 25€ for study participation and could receive an additional performance-dependent bonus in the range of 0–17€.

Stimulus material and experimental design

All stimuli were presented using MATLAB (version 8; MathWorks, Natick, MA) and the Psychophysics Toolbox extension (version 3; Brainard, 1997).

Main decision task

We decided to use a version of the random dot task, because we found it to be a well-established paradigm in the perceptual decision-making literature to elicit motion-related activation in Area V5 (e.g. Gold and Shadlen, 2007; Gallivan *et al.*, 2018). The task was designed to require the computation and integration of two distinct attribute values. For this purpose, stimuli varying in the attribute dimensions color and motion direction were employed. Each stimulus consisted of a set of 200 dots presented within a circular aperture in front of a black background (dot radius 0.07° of visual angle, aperture radius 2°, Figure 1A). Across trials, color and motion of the stimulus varied with respect to six different levels (color: blue, red, turquoise, green, brown and pink; motion direction: upward, downward, up-left, up-right, down-left and down-right with a uniform angle of 60° between directions). During each trial, all dots were constantly displayed in the same color and moved coherently into one direction (dot velocity 4°/s). Dots reaching the aperture limit were randomly replotted at the opposite semicircle (orthogonal to the current motion direction) according to a beta function ($\alpha = 1.9$, $\beta = 1.9$) to maintain an even density distribution of dots within the circular display. For both color and motion, each of the six attribute levels was associated with a monetary value taken from the set [–0.15, –0.10, –0.05, 0.05, 0.10, 0.15€]. These associations between particular monetary values and attribute levels were counterbalanced across participants and had to be acquired in a separate learning session (described in the section "learning task").

In each trial, participants had to identify the attribute values of the current color and motion direction, while being required to maintain ocular fixation throughout the trial at a cross displayed in the center of the circular aperture. These values then had to be integrated, as the sum of attribute values indicated the overall value of the stimulus. Stimuli covered all possible combinations of attribute values except those summing to 0€ (i.e. $6 \times 6 = 36$ unique combinations, out of which the 30 combinations without zero-sums were included in the experiment). Note that motion and color value are orthogonal to each other in this design, which allows for an independent assessment of their respective effects. Based on the integrated overall value, participants had to decide whether to accept or reject the current offer (i.e. optimal choices result in accepting all positive overall values and rejecting all negative ones). Participants indicated their choices by pressing one of two designated buttons on a response box using their right index/middle finger (accept/reject), which terminated

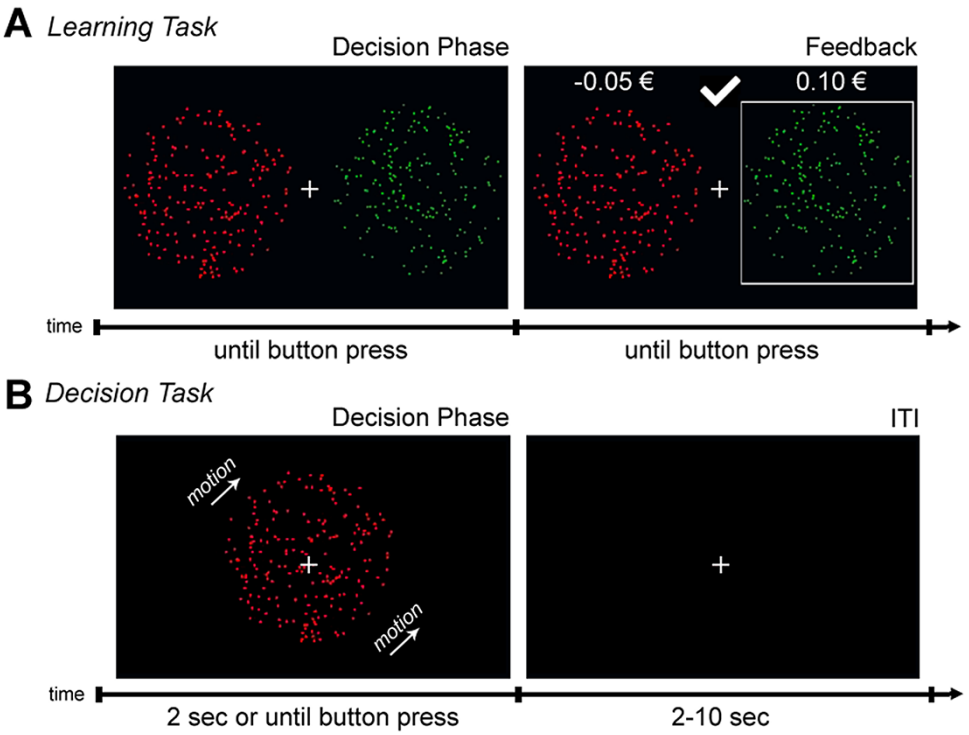

Fig. 1. Experimental design. A) Learning task. Participants had to indicate whether the left or right dot field represented a higher value and received feedback on the values of both attributes and their choice accuracy. The figure shows an example of a color value trial in which dots within both circular apertures were static but varied in their constituent color attribute. Motion value trials (which are not displayed) were designed analogously, except that dots were uniformly displayed in gray and varied with respect to their motion direction. B) Main decision task. In contrast to the learning task, participants were presented with a single circular aperture, within which the dots varied on a trial-by-trial basis with regard to both attributes. The subjects' task was to indicate whether they want to accept or reject a stimulus based on the sum of both attribute values. The decision screen terminated either by button press responses or after reaching a time limit of 2 s.

the current trial. If no response was given within 2 s after stimulus onset, the trial was automatically terminated and classified as if the inferior option had been chosen (i.e. stimuli with negative overall values were counted as accepted, and those with positive ones as rejected). Duration of inter-trial intervals (during which only the fixation cross remained on the screen) was randomized between 2 and 10 s according to a truncated exponential function ($\lambda = 6$, mean sec 5.1 ± 2.2 s.d.; Dale, 1999). After completion of the experiment, the overall values chosen for each trial were summed up and paid out as a monetary bonus (possible range: 0–17€) in addition to the monetary compensation for participation.

Our design makes use of a categorical manipulation of attribute levels. A dimensional manipulation would have also been possible (i.e. a spectrum from weak to strong motion opposed to six different motion directions), but we hypothesized that it might add more uncertainty to our task, because it can be more difficult for participants to assess the precise value of a stimulus in a dimensional design. By reducing this level of uncertainty, we reasoned that we can be more confident that participants made their choices with the intended attribute values and reduce noise in our data.

Learning task

To establish the associations between each of the individual attribute levels (i.e. the particular colors and motion directions) to one of the six monetary values, subjects completed an offline learning task (Figure 1B) in the days before scanning (mean days 2 ± 0.4 s.d.). Monetary values for color and motion levels were learned in separate blocks. During motion blocks, participants were presented with two apertures (4.5° to the left and right of the central fixation cross) each composed of 200 moving dots drawn in gray (dot and aperture size identical to the main decision task). For each trial, motion was coherent within each aperture (one of the six directions described above), but simultaneously presented apertures never showed the same direction. The participants' task was to indicate whether the left (index finger) or right (middle finger) dot field embodied the higher monetary value. After button presses, participants received feedback whether their choice was correct and corresponding monetary values were displayed above both apertures. Trials of color blocks were designed in an analogous manner with the difference that dots within each aperture remained stationary, but varied in their constituent color. Participants did not practice on stimuli combining both attributes in the learning task, to make sure that the combined experimental stimuli are not over-learned and attributes need to be actively integrated in the main decision task.

Each block consisted of 30 trials, which included two occurrences of all possible attribute value combinations. Participants completed a minimum of six blocks, including three motion and color blocks arranged in an alternating order (with the starting block type being counterbalanced across participants). After the sixth block, the task ended if participants achieved an accuracy of at least 95% during the last two blocks. If the accuracy criterion was not achieved, participants had to complete another motion and color block until it was satisfied.

Scanning session

Repetition of learning task

On the day of scanning, participants first repeated one motion and color block of the learning task outside of the scanner with 15 trials per block.

Decision task

After the learning task repetition, participants were placed in the MRI scanner. Before the recording of the first run, participants completed 15 practice trials (randomly taken from the available stimulus set) after each of which trial-wise feedback on their earnings and the values of presented attributes were displayed. The subsequent main decision task was separated into five runs with 60 trials each, with all of the 30 unique attribute value combinations (see experimental design of the main decision task) occurring twice per run. As a consequence, each motion and color value type (–0.15, –0.10, –0.05, 0.05, 0.10 and 0.15€ for each attribute) was presented 50 times throughout the entire experiment. Trial ordering was fully randomized with the exception that unique attribute value combinations were not allowed to occur twice within the first 30 trials of a run. There was no trial-wise feedback on participants' performance during the main decision task, but the total amount of earnings was displayed during brief pauses between runs.

Localizers

After completion of the decision task, motion and color localizers were conducted (order counterbalanced between participants). Both localizer tasks consisted of ten 24-s trials separated by a 12-s inter-trial interval. During the motion localizer task, subjects were presented with the same circular aperture as during the main experiment, which in contrast contained 200 dots drawn in gray. During the 24-s motion phase, dots moved coherently into a randomly chosen direction which was changed every second. During the 12-s static phase (i.e. the inter-trial interval), dots were repositioned to a new random location within the aperture after every second.

The color localizer task was designed in an analogous manner with a 24-s color phase and a 12-s inter-trial interval (achromatic phase). Stimuli consisted of a 6 × 6 checkerboard with each of the particular squares (0.4° × 0.4°) drawn in colors of random RGB values, and each square changing its color every second. The achromatic stimulus was geometrically identical, but squares changed their appearance every second only in achromatic space.

fMRI data acquisition and preprocessing

Functional data

Imaging was conducted on a 3-T Siemens Tim Trio MRI scanner (Siemens, Erlangen, Germany) with a 12-channel head coil. Functional volumes consisted of 33 continuous slices that were acquired in descending order by using a T2*-weighted gradient-echo sequence [repetition time (TR): 2 s; echo time (TE): 30 ms; matrix size: 64 × 64; field of view (FOV): 192 mm; flip angle: 78°; inter-slice gap: 0.75 mm; final voxel size: 3 × 3 × 3.75 mm]. For each participant, 133 volumes were recorded for each localizer task and an average number of 977 volumes for the main decision task (dependent on reaction times). To allow for steady-state magnetization, two dummy scans were acquired at the beginning of each run and discarded.

Structural data

For registration purposes, a high-resolution, T1-weighted structural volume was acquired from every subject after completion of the decision and localizer tasks using a magnetization-prepared rapid gradient-echo (MPRAGE) sequence (192 slices; TR: 1900 ms; TE 2.52 ms, matrix size: 256 × 256; FOV: 256 mm; flip angle: 9°; final voxel size: 1 × 1 × 1 mm).

fMRI data analysis

Preprocessing

Preprocessing was performed using SPM12 (Wellcome Trust Centre for Neuroimaging, UCL, London) and MATLAB (version 8; MathWorks, Natick, MA). Functional images were realigned, slice-time corrected, spatially normalized to the template of the Montreal Neurological Institute (MNI) and smoothed using a Gaussian kernel of 8-mm full-width at half-maximum.

GLM analysis of decision task

The decision task data were analyzed by means of two different general linear models (GLMs) for each participant. For both GLMs, event regressors were constructed as boxcar functions beginning at the time of stimulus onsets and the duration of the respective choice period.

For GLM1, regressor R1 comprised all trials during which participants made correct choices (accepting positive and rejecting negative overall values). Five linear parametric modulators of regressor R1 were included in the model to analyze neural correlates of the following decision variables: P1) motion value, P2) color value, P3) motion salience (absolute motion value), P4) color salience (absolute color value), and P5) absolute difference between motion and color value. The latter parametric modulator was included to investigate comparator regions, which could be responsible for the integration of the two attribute values. Note that P5 is not significantly correlated with P1 ($r = 0$), P2 ($r = 0$), P3 ($r = 0.09$) or P4 ($r = 0.09$). All variables were z-transformed before they were added to the model. In addition, to minimize the error term of GLM1, an additional regressor R2 comprising all incorrect choice trials was included, as well as six movement regressors R3–R8 from the realignment procedure.

GLM2 was created to analyze overall value and overall salience. These variables were not included in GLM1 due to multicollinearity which would result from significant correlations between overall values and attribute values (for each attribute $r = 0.77$, $P < 0.001$), and between overall salience and attribute salience ($r = 0.3$, $P < 0.001$). GLM2 was designed in an analogous manner to GLM1, but only included two instead of four parametric modulators: P1) overall value (sum of motion and color value) and P2) overall salience (absolute overall value).

For both models, all regressors were convolved with the canonical hemodynamic response function (HRF) and regressed against the blood-oxygen-level-dependent (BOLD) signal in each voxel. Parametric modulators were not orthogonalized to each other, allowing regressors to fully compete for explained variance. First-level contrasts were constructed by weighting all parametric modulators over baseline and submitted to second-level random-effects group analyses for statistical analysis. All statistical parametric maps from group analyses were thresholded at $P < 0.001$ (uncorrected) for voxel-level inference with a minimum cluster-size criterion of 15 contiguous voxels, and subsequent cluster-level family-wise error (FWE) correction for multiple testing at $P < 0.05$.

GLM analysis of localizer tasks

The GLMs for analyses of motion and color localizers included the following regressors: R1) boxcar function for the motion/color phase, R2) boxcar function for the static/achromatic phase and R3–R8) movement regressors as covariates of no interest. Again, regressors were convolved with the canonical HRF and regressed against the BOLD signal in each voxel. First-level contrasts were constructed by separately weighting R1 and R2 over baseline, as well as R1 > R2. These contrasts were subsequently submitted to second-level random-effects group analyses (paired *t*-tests) for statistical evaluation.

ROI analyses

To investigate whether attribute values are systematically represented in Areas V4 and V5, a region of interest (ROI) analysis was performed in two steps. First, regions maximally responsive to functional localizers were identified. To this end, 9-mm spheres were centered at peak activations in left and right V5 derived from the group analysis of the motion localizer, and at peak activations in left and right V4 derived from the group analysis of the color localizer. Within each of these four spheres, peak activations of the respective localizer were identified for each individual participant, and 6-mm spheres were centered at these coordinates. Second, these individually adapted spherical ROIs were used to extract mean beta weights from parametric modulators for motion and color value (P1 and P2 of GLM1) for each participant. The extracted beta values were then analyzed for significance via a repeated-measures analysis of variance (rm-ANOVA) with the factors attribute value (motion/color), region (V5/V4) and hemisphere (left/right). Accordingly, the two-way interaction between attribute value and region indicates whether the respective attribute values are systematically represented in V5 and V4, and the three-way interaction between all factors further allows testing for a hemisphere-specific effect. The remaining interactions (hemisphere × region, hemisphere × attribute value) were included in the analysis, but were of no interest to our research questions.

Results

Behavioral results

In the learning task, the mean accuracy for learning color values (92.8 ± 3.8% s.d.) was higher than the mean accuracy for motion values (85.9 ± 6.6% s.d.; paired *t*-test, $t_{24} = -5.57$, $P < 0.001$) across all blocks. However, all participants successfully completed the

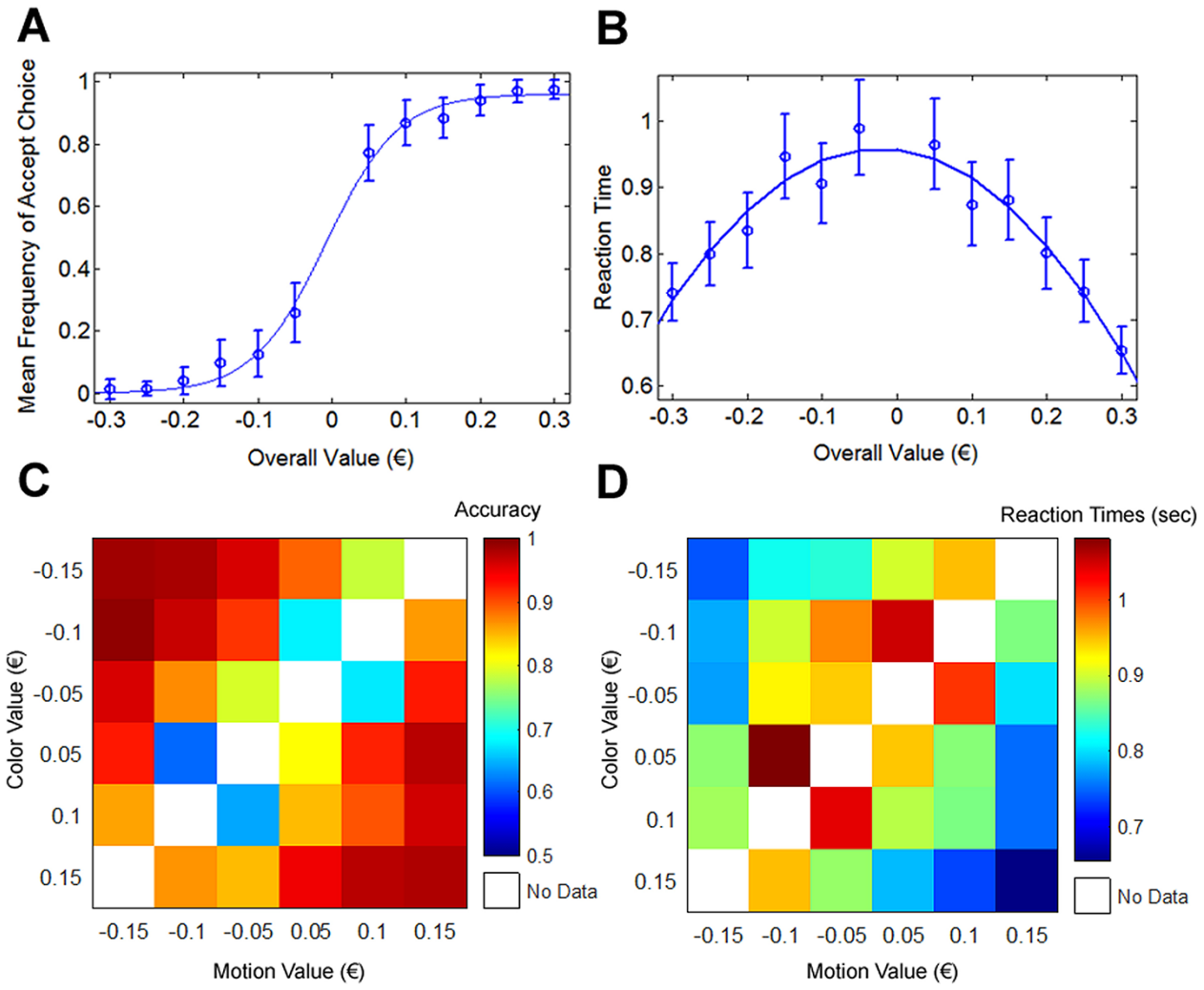

Fig. 2. Behavioral results. A) Mean frequency of accept choices plotted against overall values (fitted with a sigmoid function). Error bars represent the standard error of the mean. B) Mean reaction time plotted against overall values (fitted with a quadratic function). C) Mean accuracy for each combination of attribute value types. D) Mean reaction time for each combination of attribute value types.

learning task and achieved an accuracy of 98.1 (s.d. 2.5%) for both attributes during their last two blocks (mean of additional block number 4.4 ± 4 s.d.).

Participants showed a high level of accuracy in the main decision task (mean 87.5 ± 6.8% s.d.; one-sample *t*-test against chance level, $t_{24} = 27.56$, $P < 0.001$; Figure 2A and C) and responded on average 873 ± 157 ms (s.d.) after trial onset. Single-subject multiple linear regression models were used to estimate the effects of the overall value and the absolute attribute value difference of a stimulus (which are orthogonal to each other) on reaction times (Figure 2B and D). This analysis was performed to test whether overall value or attribute value similarity would facilitate choices. Regression coefficients showed a negative effect for overall value (one-sample *t*-test, $t_{24} = -4.27$, $P < 0.001$) and a positive effect for absolute attribute values difference ($t_{24} = 3.87$, $P < 0.001$), suggesting that participants were able to respond faster for stimuli with high overall values and high attribute value similarity. Furthermore, the effect of overall value and absolute attribute value difference on decision accuracy was tested using logistic regression models. In this analysis, absolute attribute value difference had a significant negative influence on accuracy (one-sample *t*-test, $t_{24} = -4.34$, $P < 0.001$), suggesting that participants were more accurate when attribute values were similar, but the effect of overall value was not significant ($t_{24} = -0.92$, $P = 0.367$).

Additional regression models were used to analyze differences in the processing of motion and color values. These variables were not included in the regression models above, since they are highly correlated with overall value ($r = 0.77$) and the absolute attribute value difference ($r = 1$), respectively. In a logistic regression model predicting task accuracy, neither motion (one-sample *t*-test, $t_{24} = -0.38$, $P = 0.704$) nor color value ($t_{24} = -0.42$, $P = 0.677$) were significant. In a linear regression model predicting reaction times, motion value was not significant ($t_{24} = -1.86$, $P = 0.076$), but color value had a significant negative impact ($t_{24} = -2.80$, $P = 0.009$). However, a paired *t*-test between the regression coefficients of motion and color value did not reveal a significant difference ($t_{24} = 0.65$, $P = 0.522$). This suggests that, on average, participants paid approximately equal attention to both attributes.

fMRI results

Value

The first goal of the fMRI analysis was to identify regions that are involved in valuation processes. Using GLM1, we were able to identify those regions that are specifically involved in the valuation of individual stimulus attributes (i.e. the particular values assigned to the stimulus' motion and color). We only report positive parametric modulations, because significant negative modulations were not observed. For motion value, activity during correct decision trials showed a significant positive parametric modulation in regions including PCC and left posterior inferior temporal gyrus (PIT; Figure 3A, Table 1), whereas a significant positive parametric modulation by color value was observed in ventral striatum and PCC (anterior to the PCC cluster for motion value; Figure 3B, Table 1). However, a direct comparison of motion- and color-related parametric effects using paired *t*-tests did not reveal significant differences at the whole-brain level. To further explore this relationship, the aforementioned regions were used as *post hoc* ROIs (PIT and PCC cluster of the motion value contrast, and ventral striatum and PCC cluster of the color value contrast, thresholded at $P_{unc} = 0.001$) and mean beta values for motion and color value within these ROIs were compared via paired *t*-tests (Bonferroni-corrected *P*-value criterion of 0.05/4 = 0.0125). In line with the whole-brain results, the analysis did not reveal significant differences between motion- and color-related parametric effects [PIT: $t(24) = 1.5$, $P = 0.15$; PCC_{motion}: $t(24) = 0.4$, $P = 0.71$; PCC_{color}: $t(24) = -2.4$, $P = 0.02$; ventral striatum: $t(24) = -0.7$, $P = 0.5$]. While these results do not ultimately disprove the existence of attribute-specific valuation, they nevertheless indicate that motion and color value computations do not seem to recruit clearly separable, attribute-specific valuation modules within the current study.

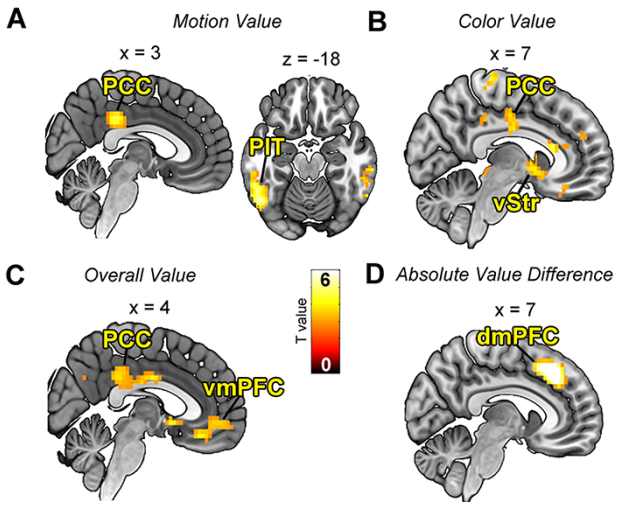

Fig. 3. Brain regions showing significant activations at the group level for A) motion value (GLM1), B) color value (GLM1), C) overall value (GLM2) and D) absolute attribute value differences (GLM1). For illustration purposes, *t*-maps (from second-level one-sample *t*-tests on parameter estimates of respective parametric modulators) are thresholded at $P_{unc} < 0.001$ with a cluster extent threshold of $k_E = 15$. Labeled clusters survive cluster-level FWE correction at $P_{FWE} < 0.05$.

Beyond the analyses of attribute valuation, GLM2 allowed to investigate which neural regions take part in the computation of overall value (i.e. the integrated value in terms of summed attribute values). We observed clusters showing a significant positive modulation of task-related activity by the stimulus' overall value in regions including left dorsolateral prefrontal cortex (dlPFC) and vmPFC (Figure 3C, Table 2). Compared to the analyses of individual attribute values, the results reveal partially overlapping neural regions (such as PCC and left PIT), which is to be expected based on the intrinsic correlation between overall and attribute-specific values. Due to this correlation, it cannot be directly assessed in a statistically valid way whether processes of overall and attribute-specific valuation show systematic neural differences. However, on the descriptive level, a significant cluster in vmPFC was only observed for parametric modulation of the stimulus' overall value, whereas this cluster was not significant in analyses of attribute-specific valuation. This pattern fits well to previous studies which suggested that vmPFC integrates information from multiple sources of evidence to an overall value (Hare *et al.*, 2009; Padoa-Schioppa and Cai, 2011; Levy and Glimcher, 2012; Rangel and Clithero, 2013).

Absolute attribute value difference

The absolute difference between motion and color values was used as a variable in GLM1 to identify comparator regions that estimate differences between attribute values (Basten *et al.*, 2010;

Table 1. Brain regions showing task-related activation in GLM1. Height threshold: $P_{unc} < 0.001$, $T_{24} = 3.47$. Extent threshold: $k_E = 15$ voxels. All activations survive whole-brain correction for multiple comparisons at the cluster level ($P_{FWE} < 0.05$). Abbreviation: MOG, middle occipital gyrus

Region	Side	MNI coordinates			k_E	T_{max}	P_{FWE} (cluster level)
		x	y	z			
Motion value							
Posterior inferior temporal gyrus	L	−51	−61	−18	171	5.74	0.003
Posterior cingulate cortex		0	−31	35	105	5.03	0.025
Superior parietal lobe	L	−39	−76	46	157	4.68	0.005
Color value							
Ventral striatum		−9	14	−3	127	6.17	0.007
Superior parietal lobe	R	18	−58	65	90	5.74	0.028
Posterior/middle cingulate cortex		−15	−22	35	171	5.49	0.002
Superior frontal sulcus	R	27	11	43	103	5.48	0.017
Anterior cingulate cortex		9	23	16	85	5.37	0.034
Dorsolateral prefrontal cortex	L	−48	35	16	167	3.91	0.002
Motion salience							
Posterior cingulate cortex		−9	−46	31	240	9.02	<0.001
LG/TPJ/MOG	R	21	−58	−14	1670	8.56	<0.001
Posterior/middle cingulate cortex		−3	−16	39	373	6.92	<0.001
Superior temporal gyrus/IPL/TPJ/mid-insular cortex/vmPFC	L	−54	−46	20	1504	5.87	<0.001
Superior frontal gyrus	L	−18	50	35	495	5.80	<0.001
Inferior frontal gyrus	R	48	38	1	76	5.75	0.022
Mid-insular cortex	R	33	−1	1	133	4.93	0.002
Superior frontal gyrus	R	15	41	46	90	4.50	0.011
Color salience							
IPL/middle temporal gyrus	L	−54	−13	−29	1610	9.08	<0.001
IPL/middle temporal gyrus	R	63	−52	20	1466	7.40	<0.001
Posterior cingulate cortex		−12	−49	31	322	6.59	<0.001
Inferior frontal gyrus	R	45	38	−10	84	6.50	0.009
Superior frontal gyrus/vmPFC	L	−18	29	58	345	6.06	<0.001
Posterior cingulate cortex		0	−16	39	215	5.91	<0.001
Hippocampus/putamen	L	−21	−4	5	262	5.67	<0.001
Fusiform gyrus	R	42	−55	−18	189	5.60	<0.001
Cerebellum	R	21	−82	−33	84	5.35	0.009
Inferior frontal gyrus	L	−51	32	−14	70	5.27	0.02
Middle frontal gyrus	L	−33	29	46	111	5.05	0.002
Absolute attribute value difference							
Dorsomedial prefrontal cortex		0	32	43	287	8.85	<0.001
Inferior frontal gyrus	R	33	26	−6	97	7.78	0.012
Middle frontal gyrus	R	45	14	43	114	5.89	0.006
Inferior frontal gyrus	L	−33	20	1	72	5.75	0.038
Superior frontal gyrus	R	27	17	58	92	5.66	0.015
IPL	R	48	−52	58	113	5.56	0.006

Philiastides *et al.*, 2010). By this means, we observed a significant positive modulation of task-related hemodynamic activity within the dorsomedial prefrontal cortex (dmPFC; Figure 3D, Table 1).

Salience

A significant positive modulation of task-related activity by motion salience (GLM1) was observed in regions including bilateral temporoparietal junction (TPJ), right lingual gyrus (LG) and PCC (Figure 4A, Table 1), whereas a positive modulation by color salience (GLM1) was found in bilateral inferior parietal lobe (IPL), bilateral anterior temporal cortex (AT) and PCC (see Figure 4B and Table 1). Direct comparisons of motion and color salience effects by paired *t*-tests revealed no activations surviving our significance criterion, suggesting that neither attribute salience had a significantly stronger effect nor relies on specialized processing modules in our current decision task. For modulation by overall stimulus salience (GLM2), partially overlapping regions including bilateral TPJ and PIT (Table 2) were observed.

Attribute processing in V5/V4

Motion and color localizers were conducted to identify regions specifically involved in the processing of physical stimulus attributes (i.e. motion and color), independent of valuation processes (Figure 5A and B). As expected, the motion localizer revealed significant activation in the group analysis within bilateral V5 (left V5: $T_{24} = 5.89$, $k = 107$, cluster-level $P_{FWE} = 0.036$, $x = -48$, $y = -67$, $z = 5$; right V5: $T_{24} = 6.28$, $k = 109$, cluster-level $P_{FWE} = 0.034$, $x = 42$, $y = -64$, $z = 9$), and the color localizer within bilateral V4 (left V4: $T_{24} = 8.66$, $k = 753$, cluster-level $P_{FWE} < 0.001$, $x = -30$, $y = -73$, $z = -14$; right V4: $T_{24} = 8.63$, $k = 753$, cluster-level $P_{FWE} < 0.001$, $x = 30$, $y = -73$, $z = -14$).

ROI analyses were performed to test whether regions specialized in processing the physical attributes of the stimuli (i.e. V5 for motion and V4 for color) also compute the respective attribute values. For this purpose, beta estimates from parametric modulations by motion and color value were extracted from V5 and V4 in both hemispheres, and entered into an rm-ANOVA with factors attribute value (motion/color), region (V5/V4) and hemisphere (left/right). Neither the two-way interaction between

Table 2. Brain regions showing task-related activation in GLM2. Height threshold: $P_{unc} < 0.001$, $T_{24} = 3.47$. Extent threshold: $k_E = 15$ voxels. All activations survive whole-brain correction for multiple comparisons at the cluster level ($P_{FWE} < 0.05$)

		MNI coordinates					
Region	**Side**	**x**	**y**	**z**	k_E	T_{max}	P_{FWE} **(cluster level)**
Overall value							
Dorsolateral prefrontal cortex	L	−51	29	24	136	6.12	0.023
Superior parietal lobe	L	−27	−76	46	187	5.30	0.007
Posterior inferior temporal gyrus	R	60	−25	−25	216	5.27	0.004
Posterior cingulate cortex		−21	−22	35	499	5.14	<0.001
Inferior temporal gyrus	L	−54	−58	−18	215	4.97	0.004
Ventromedial prefrontal cortex		0	35	−18	212	4.74	0.004
Overall salience							
Inferior temporal gyrus/TPJ	R	48	−34	1	1607	6.73	<0.001
TPJ/postcentral gyrus	L	−51	−22	35	415	6.11	<0.001
Inferior temporal gyrus	L	−33	−46	−18	181	5.87	<0.001
Occipital lobe	L	−33	−88	1	302	5.46	<0.001
Rostral anterior cingulate cortex		18	32	−10	108	5.24	0.007
Hippocampus	R	30	−10	−18	96	4.92	0.012
Hippocampus	L	−27	−16	−14	81	4.53	0.024
Middle temporal gyrus	L	−39	−64	5	69	4.11	0.042

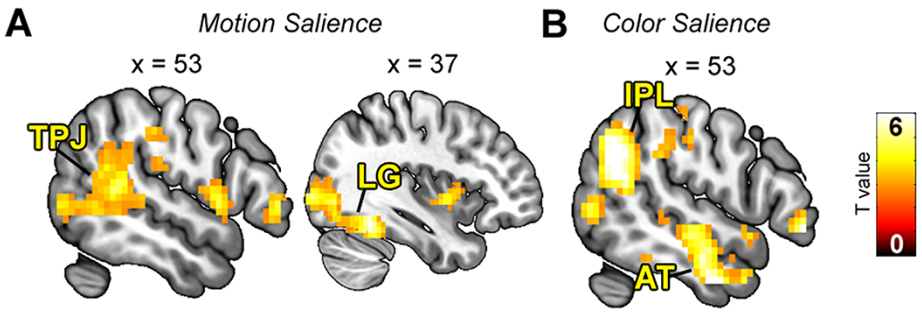

Fig. 4. Brain regions showing significant group-level activations for A) motion salience and B) color salience (GLM1). For illustration purposes, t-maps (from second-level one-sample t-tests on parameter estimates of the respective parametric modulator) are thresholded at $P_{unc} < 0.001$ with a cluster extent threshold of $k_E = 15$. Labeled clusters survive cluster-level FWE correction at $P_{FWE} < 0.05$.

attribute value and region [$F(1, 24) = 0.51$, $P = 0.48$] nor the three-way interaction between attribute value, region and hemisphere [$F(1, 24) = 0.53$, $P = 0.47$] were significant, which does not support the hypothesis that attribute values are systematically processed in V5 and V4. Furthermore, there were no significant effects for the main effects or the remaining interactions of no interest [attribute value: $F(1, 24) = 0.01$, $P = 0.93$; region: $F(1, 24) = 3.15$, $P = 0.09$; hemisphere: $F(1, 24) = 0.12$, $P = 0.73$; hemisphere × region: $F(1, 24) = 0.51$, $P = 0.48$; hemisphere × attribute value: $F(1, 24) = 2.57$, $P = 0.12$].

Discussion

The current experiment investigated how attribute values are processed and integrated in the brain during decision-making. In particular, we tested the competing hypotheses whether (i) distinct regions that specialize in the processing of a particular physical attribute also compute the respective attribute value (Philiastides *et al.*, 2010; Lim *et al.*, 2013), or whether (ii) attribute values are collectively processed in a general valuation network consisting of vmPFC, PCC and ventral striatum (Levy and Glimcher, 2012; Bartra *et al.*, 2013; Clithero and Rangel, 2014). To differentiate between these hypotheses, we used a choice task in which monetary values were associated with the attributes motion and color, whose physical properties are known to be processed in specialized brain regions, namely Area V5 for motion (Watson *et al.*, 1993) and Area V4 for color (McKeefry and Zeki, 1997).

Whole-brain analyses showed that activity in PCC and ventral striatum correlated with color value, whereas activity related to motion value occurred in PCC and left PIT. In a direct comparison, we did not detect any region that had a specifically stronger representation of one compared to the other attribute value. This lack of specificity suggests that the computation of particular attribute values is not realized within specialized cortical modules, but is instead accomplished in a dynamic manner within a network comprising PCC, PIT and ventral striatum. Consistent with this idea, ROI analyses did not reveal a systematic representation of motion value in V5 and color value in V4, which does not support the hypothesis that attribute values and physical properties of attributes are computed in the same regions. Taken together, our data thus provide concordant evidence for the hypothesis that values are homogeneously processed within a general valuation network.

In contrast to our results, previous studies have supported the hypothesis that attribute values are computed in attribute-specific regions. In an experiment by Lim *et al.* (2013), participants had to evaluate t-shirts based on how much they liked both the appearance and meaning of Korean symbols that were printed on them. The authors found that activity in fusiform gyrus correlated with visual values, whereas activity in superior temporal gyrus correlated with semantic values. Furthermore, a study by

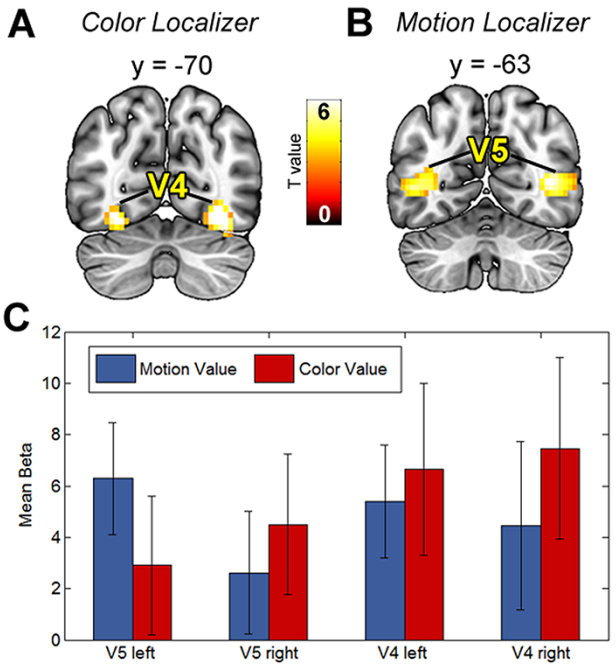

Fig. 5. Activations of localizer tasks in A) bilateral V4 (color localizer) and B) bilateral V5 (motion localizer). For illustration purposes, *t*-maps are thresholded at $P_{unc} < 0.001$ with a cluster extent threshold of $k_E = 15$. All clusters survive cluster-level FWE correction at $P_{FWE} < 0.05$. C) Mean beta estimates of parametric modulation by motion and color value in bilateral V5 and V4 (regions adapted to single-subject peaks of localizers; for details see Methods section). As the corresponding results of the rm-ANOVA with factors attribute value (motion/color), region (V5/V4) and hemisphere (left/right) indicate, the data do not reveal a systematic representation of motion and color value in V5 and V4, respectively.

Philiastides *et al.* (2010) demonstrated in a probabilistic choice task that activity in fusiform face area correlates with the value of face stimuli, and parahippocampal place area with the value of house stimuli. For both studies, the brain regions correlating with attribute values have also been associated with the processing of physical attribute properties, which is not confirmed by data of the present experiment.

This discrepancy could originate from several factors. First, our design allowed us to differentiate value- and salience-related effects for each attribute, which was not possible in the experiment by Philiastides *et al.* (2010). As such, value correlations in fusiform face area and parahippocampal place area might originate from differences in salience instead of value. Second, the motion and color attributes in our paradigm are robustly associated with well-defined regions in Areas V5 and V4. Arguably, this connection is less straightforward in the aforementioned studies, in particular for the semantic attribute of the study by Lim *et al.* (2013), because the effects are more distributed across brain regions. Therefore, it is less clear whether these results actually arise from attribute-specific regions, as it is more difficult to precisely determine ROIs. Third, we used abstract, novel stimuli in our paradigm, whereas the other studies used more common stimuli (faces, houses and t-shirts). A possible mechanism for neural attribute valuation could be that novel stimuli are first computed in a domain-general network, in which attribute values are processed homogeneously, but if stimuli are more familiar and need to be evaluated frequently, the computation of attribute values shifts toward the respective attribute-specific regions. Thus, the processing of attribute values might change as a function of learning, which could reconcile the effects of our experiment with previous studies.

Apart from analyses of single attribute values, we also investigated brain activity related to the overall value of stimuli and found correlations with activity in, among other regions, left dlPFC and vmPFC. Due to the correlation between attribute values and overall values, our results show an overlap in their effects on the neural level, which cannot be easily disentangled on statistical grounds. Nonetheless, it is of note that whole-brain analyses of overall values revealed a significant correlation for a cluster in vmPFC, which was not found in our analyses of attribute values (neither for motion nor for color value). This is in line with the view that vmPFC is one of the main regions that represent integrated value signals, which previous studies have argued for (Hare *et al.*, 2009; Padoa-Schioppa and Cai, 2011; Levy and Glimcher, 2012; Rangel and Clithero, 2013; Pelletier and Fellows, 2019). From this perspective, decision problems are deconstructed into subproblems (such as the computation of single attribute values) that are resolved in distributed brain regions, and the available evidence is then ultimately combined to a unified value signal in vmPFC, which is consistent with our findings.

Furthermore, we analyzed brain activity related to the absolute difference between simultaneously presented attribute values. This variable could be an indicator for regions that compare attribute values and thus estimate difference signals that drive the process of value integration. Consistent with findings from other studies (Basten *et al.*, 2010; Hare *et al.*, 2011), we found that absolute differences between attribute values were mainly associated with activity in dmPFC. Moreover, the behavioral data revealed that larger differences between attribute values were associated with longer reaction times and a lower choice accuracy, which suggests that these trials were more difficult for participants. One explanation for this relationship could be that larger attribute value differences induce a higher need for value integration and thus demand more cognitive resources. When attribute value differences are small, each attribute value in isolation already provides a good estimate for a stimulus' overall value, and value integration is less important for effective decision-making. But when attribute value differences are large, it is crucial to integrate the underlying attribute values into a representative overall value to make optimal decisions. Thus, dmPFC could be responsible for the estimation of attribute value difference signals that indicate the need for value integration and form the basis for the computation of overall values.

In addition to analyses of value-related effects, we also explored salience-related effects. Salience refers to the subjective importance of a stimulus (Maunsell, 2004; Zink *et al.*, 2006; Litt *et al.*, 2011; Leathers and Olson, 2012; Kahnt and Tobler, 2013; Kahnt *et al.*, 2014) that guides attention and prioritizes the processing of particular stimuli over less important ones. In the context of this experiment, both high positive as well as high negative attribute values have a high salience, because both have a large impact on decisions. The more salient an attribute value is (i.e. the higher the absolute attribute value is), the less likely it is that the other attribute value will outweigh its influence. Hence, when time is limited, it is efficient to selectively process attributes with higher salience and pay less attention to attributes with lower salience (Kahnt and Tobler, 2013). We investigated brain regions that could realize such a selection mechanism by analyzing neural correlates of absolute attribute values for motion and color. As a result, motion salience was associated with bilateral TPJ, right LG and PCC, whereas bilateral IPL, AT and PCC were correlated with color salience. This network could therefore be

responsible for allocating attentional resources and assigning priority to the attribute that is most relevant in a given situation. Notably, activity in PCC was related to attribute value, attribute salience as well as overall value, which suggests that it plays a central part for neural information processing in value-based decision-making.

There are some limitations that have to be taken into account in the interpretation of our findings. First, our color and motion attributes have different perceptual properties, but they were both encoded via monetary values. However, in real-world decisions, the attributes that have to be combined often have different types of value encoding. For example, in the evaluation of a car, the price has a monetary value encoding, the design an aesthetic value encoding, and the safety rating is encoded as risk to our health. Arguably, the attributes in our experiment still qualify as different attributes, in the same way that the interior and exterior design of a car can be evaluated independently, even though both have an aesthetic value encoding. But it is an open question in what way our results would differ if the attributes had different types of value encodings. Second, while we did not observe representations of attribute values in V4 or V5, we cannot decisively rule out that possibility, since strong conclusions from null findings are not statistically justified. Third, we modeled events in our GLM for the duration of the whole choice period, but it is possible that explicitly modeling different stages of the decision-making process (such as attribute value identification and attribute integration) could reveal stronger and more precise effects. There is also evidence that decision attributes are processed at different rates (e.g. Sullivan *et al.*, 2014), which could have played a role in our task as well. For example, while color can, in principle, be perceived immediately, motion perception requires the observation of visual frames during a longer time span. The low temporal resolution of fMRI makes it challenging to better incorporate factors like these, but a dedicated experimental design and methods like the combination of electroencephalography and fMRI could help to disentangle subprocesses of decision-making in more detail.

So far, most studies on attribute integration in value-based decision-making (Hare *et al.*, 2009; Basten *et al.*, 2010; Philiastides *et al.*, 2010; Kahnt *et al.*, 2011; Park *et al.*, 2011; Lim *et al.*, 2013; Hutcherson *et al.*, 2015) have argued for a feed-forward model, in which attribute values are separately computed in dedicated regions, and only afterward integrated in a unifying region like vmPFC. However, some studies have argued for a more flexible model (Hunt *et al.*, 2014; Siegel *et al.*, 2015). In this view, attribute values are determined in a dynamic process that includes continuous feed-forward and feedback projections as well as competitive inhibition between attributes. Hence, the model proposes that attribute values are not computed sequentially and in isolation. Instead, there is a constant exchange of information in which value predictions are continuously updated and re-evaluated, dependent on concurrent neural computations that process factors like salience, memory or affective states. Since we observed a uniform neural network for different types of attribute values, the results of our study are in line with the latter model and support the idea that attribute values are computed in an interdependent and contextualized manner.

Acknowledgements

We thank Christine Stelzel and Ilya Veer for valuable feedback during discussions, and Gabriele Bellucci for his help in data acquisition.

Funding

This work was supported by the Berlin School of Mind and Brain, Charité Berlin, Bernstein Computational Neuroscience Program of the German Federal Ministry of Education and Research (grant 01GQ1001C), the German Research Foundation within the Collaborative Research Center 'Volition and Cognitive Control: Mechanisms, Modulations, Dysfunctions' (DFG grant SFB 940/1 2014 and 2015), the Flemish Science Foundation and the European Union's Horizon 2020 Research and Innovation Program (Marie Skłodowska-Curie grant 665501).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

A.M., C.M.S., V.U.L., H.R.H. and H.W. designed the experiment. A.M. programmed the experiment, acquired the data and analyzed the results. C.M.S., V.U.L., L.M.P., D.W., H.R.H. and H.W. provided methodological and conceptual advice for the analysis and discussion of the data. The manuscript was written by A.M. and edited by all authors.

References

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**, 412–27.

Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences USA*, **107**, 21767–72.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–6.

Clithero, J., Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, **9**, 1289–302.

Dale, A.M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, **8**, 109–14.

de Berker, A.O., Kurth-Nelson, Z., Rutledge, R.B., Bestmann, S., Dolan, R.J. (2019). Computing value from quality and quantity in human decision-making. *Journal of Neuroscience*, **39**, 163–76.

Gallivan, J.P., Chapman, C.S., Wolpert, D.M., Flanagan, J.R. (2018). Decision-making in sensorimotor control. *Nature Reviews Neuroscience*, **19**, 519–34.

Gold, J.I., Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, **30**, 535–74.

Hanks, T.D., Summerfield, C. (2017). Perceptual decision making in rodents, monkeys, and humans. *Neuron*, **93**, 15–31.

Hare, T.A., Camerer, C.F., Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*, **324**, 646–8.

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 USA*, **108**, 18120–5.

Hunt, L.T., Dolan, R.J., Behrens, T.E.J. (2014). Hierarchical competitions subserving multi-attribute choice. *Nature Neuroscience*, **17**, 1613–22.

Hutcherson, C.A., Montaser-Kouhsari, L., Woodward, J., Rangel, A. (2015). Emotional and utilitarian appraisals of moral dilemmas

are encoded in separate areas and integrated in ventromedial prefrontal cortex. *Journal of Neuroscience*, **35**, 12593–605.

Kable, J.W., Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, **10**, 1625–33.

Kable, J.W., Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, **63**, 733–45.

Kahnt, T., Heinzle, J., Park, S.Q., Haynes, J.-D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *NeuroImage*, **56**, 709–15.

Kahnt, T., Park, S.Q., Haynes, J.-D., Tobler, P.N. (2014). Disentangling neural representations of value and salience in the human brain. *Proceedings of the National Academy of Sciences USA*, **111**, 5000–5.

Kahnt, T., Tobler, P.N. (2013). Salience signals in the right temporoparietal junction facilitate value-based decisions. *Journal of Neuroscience*, **33**, 863–9.

Kanwisher, N., McDermott, J., Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, **17**, 4302–11.

Leathers, M.L., Olson, C.R. (2012). In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science*, **338**, 132–5.

Levy, D.J., Glimcher, P.W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, **22**, 1027–38.

Lim, S.-L., O'Doherty, J.P., Rangel, A. (2013). Stimulus value signals in ventromedial PFC reflect the integration of attribute value signals computed in fusiform gyrus and posterior superior temporal gyrus. *Journal of Neuroscience*, **33**, 8729–41.

Litt, A., Plassmann, H., Shiv, B., Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*, **21**, 95–102.

Ludwig, V.U., Stelzel, C., Krutiak, H., *et al.* (2014). The suggestible brain: posthypnotic effects on value-based decision-making. *Social Cognitive and Affective Neuroscience*, **9**, 1281–8.

Maunsell, J.H.R. (2004). Neuronal representations of cognitive state: reward or attention? *Trends in Cognitive Sciences*, **8**, 261–5.

McKeefry, D.J., Zeki, S. (1997). The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain*, **120**, 2229–42.

O'Doherty, J.P. (2014). The problem with value. *Neuroscience and Biobehavioral Reviews*, **43**, 259–68.

Padoa-Schioppa, C., Cai, X. (2011). The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. *Annals of the New York Academy of Sciences*, **1239**, 130–7.

Park, S.Q., Kahnt, T., Rieskamp, J., Heekeren, H.R. (2011). Neurobiology of value integration: when value impacts valuation. *Journal of Neuroscience*, **31**, 9307–14.

Pelletier, G., Aridan, N., Fellows, L.K., Schonberg, T. (2021). A preferential role for ventromedial prefrontal cortex in assessing the value of the whole in multiattribute object evaluation. *Journal of Neuroscience*, **41**, 5056–68.

Pelletier, G., Fellows, L.K. (2019). A critical role for human ventromedial frontal lobe in value comparison of complex objects based on attribute configuration. *Journal of Neuroscience*, **39**, 4124–32.

Persichetti, A.S., Aguirre, G.K., Thompson-Schill, S.L. (2015). Value is in the eye of the beholder: early visual cortex codes monetary value of objects during a diverted attention task. *Journal of Cognitive Neuroscience*, **27**, 893–901.

Peters, J., Büchel, C. (2010). Neural representations of subjective reward value. *Behavioural Brain Research*, **213**, 135–41.

Philiastides, M.G., Biele, G., Heekeren, H.R. (2010). A mechanistic account of value computation in the human brain. *Proceedings of the National Academy of Sciences USA*, **107**, 9430–5.

Rangel, A., Camerer, C., Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, **9**, 545–56.

Rangel, A., Clithero, J. (2013). The computation of stimulus values in simple choice. In: Glimcher, P., Fehr, E., editors. *Neuroeconomics: Decision Making and the Brain*, 2nd edn, San Diego: Academic Press, 125–48.

Sanfey, A.G., Loewenstein, G., McClure, S.M., Cohen, J.D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in Cognitive Sciences*, **10**, 108–16.

Serences, J.T. (2008). Value-based modulations in human visual cortex. *Neuron*, **60**, 1169–81.

Siegel, M., Buschman, T.J., Miller, E.K. (2015). Cortical information flow during flexible sensorimotor decisions. *Science*, **348**, 1352–5.

Sullivan, N., Hutcherson, C., Harris, A., Rangel, A. (2014). Dietary self-control is related to the speed with which attributes of healthfulness and tastiness are processed. *Psychological Science*, **26**, 122–34.

Suzuki, S., Cross, L., O'Doherty, J.P. (2017). Elucidating the underlying components of food valuation in the human orbitofrontal cortex. *Nature Neuroscience*, **20**, 1780–6.

Vaidya, A.R., Sefranek, M., Fellows, L.K. (2018). Ventromedial frontal lobe damage alters how specific attributes are weighed in subjective valuation. *Cerebral Cortex*, **28**, 3857–67.

Watson, J.D.G., Myers, R., Frackowiak, R.S.J., *et al.* (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, **3**, 79–94.

Zhang, Z., Fanning, J., Ehrlich, D.B., Chen, W., Lee, D., Levy, I. (2017). Distributed neural representation of saliency controlled value and category during anticipation of rewards and punishments. *Nature Communications*, **8**, 1–14.

Zink, C.F., Pagnoni, G., Chappelow, J., Martin-Skurski, M., Berns, G.S. (2006). Human striatal activation reflects degree of stimulus saliency. *NeuroImage*, **29**, 977–83.