Parallel representation of context and multiple context-dependent values in ventro-medial prefrontal cortex

Nir Moneta[1,3,4*], Mona M. Garvert[1,4,5,6], Hauke R. Heekeren[2], and Nicolas W. Schuck[1,4*]

[1]Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, 14195 Berlin, Germany
[2]Department of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany
[3]Einstein Center for Neurosciences Berlin, Charité Universitätsmedizin Berlin, 10117, Berlin, Germany
[4]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, 14195 Berlin, Germany
[5]Department of Psychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany
[6]Wellcome Centre for Integrative Neuroimaging, FMRIB, University of Oxford, Oxford, United Kingdom
[*]Corresponding authors, email moneta@mpib-berlin.mpg.de and schuck@mpib-berlin.mpg.de

March 17, 2021

Abstract

Value representations in ventromedial prefrontal-cortex (vmPFC) are known to guide decisions. But how preferable available options are depends on one's current task. Goal-directed behavior, which involves changing between different task-contexts, therefore requires to know how valuable the same options will be in different contexts. We tested whether multiple task-dependent values influence behavior and asked if they are integrated into a single value representation or are co-represented in parallel within vmPFC signals. Thirty five participants alternated between tasks in which stimulus color or motion predicted rewards. Our results provide behavioral and neural evidence for co-activation of both contextually-relevant and -irrelevant values, and suggest a link between multivariate neural representations and the influence of the irrelevant context and its associated value on behavior. Importantly, current task context could be decoded from the same region, and better context-decodability was associated with stronger (relevant-)value representations. Evidence for choice conflicts was found only in the motor cortex, where the competing values are likely resolved into action.

Introduction

Decisions are always made within the context of a given task. Even a simple choice between two apples will depend on whether the task is to find a snack, for which their color might indicate the desired sweetness, or to buy ingredients for a cake, for which a crisp texture might be more crucial. In other words, the same objects can yield different outcomes under different task contexts. Context-dependent decision-making therefore requires to retrieve not only the outcomes that are associated with different objects. Rather, it is necessary to maintain separate outcome expectations for the same choice option, and to know in which task context which outcome expectation is relevant.

Computing the reward a choice will yield given a task context is at the core of decisions [e.g. 1]. In line with this idea, previous studies have shown in a variety of species that the ventromedial prefrontal cortex (vmPFC) represents this so-called expected value (EV) [2–7], and thereby plays a crucial role in determining choices [8]. It is also known that the brain's attentional control network enhances the processing of features that are relevant given the current task context [9, 10], and that this helps to shape which features influence EV representations in vmPFC [11–13]. Moreover, the vmPFC seems to also represent the EV of different features in a common currency [14, 15]; and thus is necessary for integrating the expectations from different reward predicting features of the same object [16–18]. It remains unclear however, how context-irrelevant value expectations of presented features, i.e. rewards that would be obtained in a different task-context, might affect neural representations in vmPFC.

This is particularly relevant because we often have to do more than one task within the same environment, such as shopping in the same supermarket for different purposes. Thus we have to switch between the values that are relevant in the different contexts. Moreover, the separation between tasks can often be less than perfect, which can then lead to processing of task-irrelevant aspects. In line with this idea, several studies have shown that decisions are influenced by contextually-irrelevant information, and traces of the distracting features in cortical regions responsible on task execution [19–23]. Similarly, task-irrelevant valuation has been shown to influence attentional selection [24] as well as activity in posterior parietal [25] or ventromedial prefrontal cortex [26]. This raises the possibility that vmPFC represents different value expectations that could occur in different task contexts at the same time. In the present study we therefore investigated whether the vmPFC maintains multiple task-dependent values during choice, and how these representations influence choices, interact with the encoding of the relevant task-context, and with each other.

Previous research has indeed suggested that the role of vmPFC in decision making seems not to be restricted to representing economic values. Rather, other aspects of the current task might be encoded in this region as well [27–31]. Of particular relevance, a number of investigations have indicated that vmPFC and adjacent overlapping medial orbitofrontal cortex represents the current context or task state in humans [32–35].This task state effectively encodes which features are currently relevant and thereby determines which value expectations will guide behavior. Note, however, that these value and task-state accounts do not need to be mutually exclusive, but rather might reflect multiplexed representations within the neural activity of the vmPFC/OFC [36, 37]. Conceptualizing the role of vmPFC as representing possible task states therefore bridges beyond its traditional role as controller of economic value to a more complex role of parallel representation of task-related information, EV included.

If neural activity in vmPFC goes beyond signalling a single EV by representing more complex task structure, then it suggests that the task-context is represented in addition to the values. We therefore hypothesized that vmPFC indeed simultaneously represents the task-context, as well as task-relevant and task-irrelevant values. This idea – that values and task-context co-occur and interact – also predicts that a stronger activation of the relevant task-context will enhance the representation of task-relevant values. We investigated this question using a multi-feature choice task in which different features of the same stimulus predicted different outcomes and a task-context cue modulated which feature was relevant. We hypothesized that values associated with contextually irrelevant features affect value representations in vmPFC. Moreover, we tested whether different possible EVs were integrated into a single value representation or processed in parallel. The former would support a unique role of the vmPFC for representing *only* the EV of choice, whereas the latter would indicate that the vmPFC encodes several aspects of a complex task structure, including separate value representations for the currently relevant and irrelevant task contexts.

Results

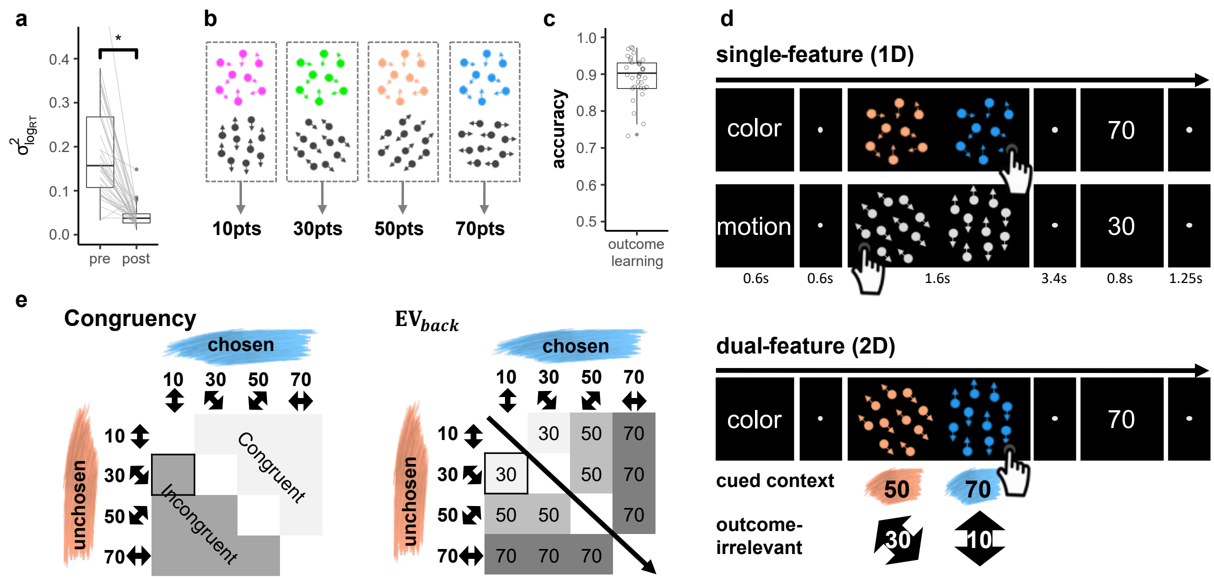


Figure 1: Task and Design a. Staircasing procedure reduced differences in detection speed between features. Depicted is the variance of reaction times (RTs) across different color and motion features (y axis). While participants' RTs were markedly different for different features before staircasing (pre), a significant reduction in RT differences was observed after the procedure (post). The staircasing procedure was performed before value learning. RT-variance was computed by summing the squared difference of each feature's RT and the general mean RT per participant. $N = 35$, $p < .001$. **b.** The task included eight features, four color and four motion directions. After the stair-casing procedure, a specific reward was assigned to each motion and each color, such that one feature from each of the contexts had the same value as it was associated with the same reward. Feature values were counterbalanced across participants. **c.** Participants were trained on feature values shown in (b) and achieved near ceiling accuracy in choosing the highest valued feature afterwards ($\mu = .89, \sigma = .06$). **d.** Single- and dual-feature trials (1D, 2D, respectively). Each trial started with a cue of the relevant context (Color or Motion, 0.6s), followed by a short fixation circle (0.6s). Participants were then presented with a choice between two clouds (1.6s). Each cloud had only one feature in 1D trials (colored dots, but random motion, *or* directed motion, but gray dots, top) and two features for 2D trials (motion *and* color, bottom). Participants were instructed to make a decision between the two clouds based on the cued context and ignore the other. Choices were followed by a fixation period (3.4s) and the value associated with the chosen cloud's feature of the cued context (0.8s). After another short fixation (1.25s) the next trial started. **e.** Variations in values irrelevant in the present task context of a 2D trial. For each feature pair (e.g. blue and orange), all possible context-irrelevant feature-combinations were included in the task, except the same feature on both sides. Congruency (left): trials were separated into those in which the irrelevant features favored the same choice as the relevant features (congruent trials), or not (incongruent trials). $\mathrm{EV_{back}}$ (right): based on this factor, the trials were characterized by different hypothetically expected values of the contextually-irrelevant features, i.e. the maximum value of both irrelevant features. Crucially, EV, $\mathrm{EV_{back}}$ and Congruency were orthogonal by design. The example trial presented in (d, bottom) is highlighted.

Behavioral results

Participants had to judge either the color or motion direction of moving dots on a screen (random dot motion kinematogramms, [e.g. 38]). Four different colors and motion directions were used. Before entering the MRI scanner, participants performed a stair-casing task in which participants had to indicate which of two shown stimuli corresponded to a previously cued feature. Motion-coherence and the speed which dots changed from grey to a target color were adjusted such that the different stimulus features could be discriminated equally fast, both within and between contexts. As intended, this led to significantly reduced differences in reaction times (RTs) between the eight stimulus features ($t_{(34)} = 7.29$, $p < .001$, Fig.1a), also when tested for each button separately ($t_{(34)} =$ Left: 6.52, Right: 7.70, ps$< .001$, Fig. S1d)

Only then, participants learned to associate each color and motion feature with a fixed number of points (10, 30, 50 or 70 points), whereby one motion direction and one color each led to the same reward (counterbalanced across participants, Fig.1b). To this end, participants had to make a choice between clouds that had only one feature-type, while the other feature type was absent or ambiguous (clouds were grey in motion clouds and moved randomly in color clouds). To encourage mapping of all features on a unitary value scale, choices in this part (and only here) also had to be made between contexts (e.g. between a green and a horizontal-moving cloud). At the end of the learning phase, participants achieved near-ceiling accuracy in choosing the cloud with the highest valued feature ($\mu = .89, \sigma = 0.06$, t-test against chance: $t_{(34)} = 41.8$, $p < .001$, Fig. 1c), also when tested separately for Color, Motion and across context ($\mu = .88, .87, .83$,$\sigma = .09, .1, .1$, t-test against chance: $t_{(34)} = 23.9, 20.4, 19.9$, ps$< .001$, respectively, Fig. S1e). Once inside the MRI scanner, one additional training block ensured changes in presentation mode did not induce feature-specific RT changes ($F_{(7,202)} = 1.06$, $p = 0.392$). These procedures made sure that participants began the main experiment inside the MRI scanner with firm knowledge of feature values; and that RT differences would not reflect perceptual differences, but could be attributed to the associated values. Additional information about the pre-scanning phase can be found in Online Methods and in Fig.S1.

During the main task, participants had to select one of two dot motion clouds. In each trial participants were first cued whether a decision should be made based on color or motion features, and then had to choose the cloud that would lead to the largest number of points. Following their choice, participants received the points corresponding to the value associated with the chosen cloud's relevant feature. To reduce complexity, the two features of the *cued task-context* always had a value difference of 20, i.e. the choices on the cued context were only between values of 10 vs. 30, 30 vs. 50 or 50 vs. 70. One third of the trials consisted of a choice between single-feature clouds of the same context (henceforth: 1D trials, Fig.1d, top). All other trials were dual-feature trials, i.e. each cloud had a color *and* a motion direction at the same time (henceforth: 2D trials, Fig.1d bottom), but only the color or motion features mattered as indicated by the cue. Thus, while 2D trials involved four features in total (two clouds with two features each), only the two color *or* two motion features were relevant for determining the outcome. The cued context stayed the same for a minimum of four and a maximum of seven trials. Importantly, for each comparison of relevant features, we varied which values were associated with the features of the *irrelevant* context, such that each relevant value was paired with all possible irrelevant values (Fig.1e). Consider, for instance, a color trial in which the color shown on the left side led to 50 points and the color on the right side led to 70 points. While motion directions in this trial did not have any impact on the outcome, they might nevertheless influence behavior. Specifically, they could favor the same side as the colors or not (Congruent vs Incongruent trials, see Fig.1e left), and have larger or smaller values compared to the color features (Fig.1e right).

We investigated the impact of these factors on RTs in correct 2D trials, where the extensive training ensured near-ceiling performance throughout the main task ($\mu = 0.91, \sigma = 0.05$, t-test against chance: $t_{(34)} = 48.48$, $p < .0001$, Fig.2a). RTs were log transformed to approximate normality and analysed using mixed effects models with nuisance regressors for choice side (left/right), time on task (trial number), differences between attentional contexts (color/motion) and number of trials since the last context switch. We used a hierarchical model comparison approach to asses the effects of (1) the objective value of the chosen option (or: EV), i.e. points associated with the features on the cued context; (2) the maximum points that could have been obtained if the irrelevant features were the relevant ones (the expected value of the background, henceforth: $\mathrm{EV}_{\mathrm{back}}$, Fig 1e left), and (3) whether the irrelevant features favored the same side as the relevant ones or not (Congruency, Fig. 1e right). Any effect of the latter two factors would indicate that outcome associations that were irrelevant in the current context nevertheless influence behavior, and therefore could be represented in vmPFC.

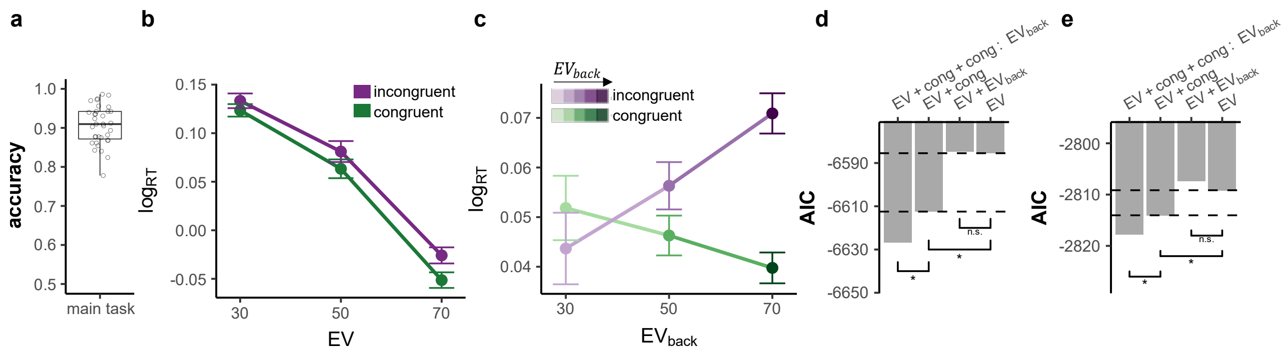


Figure 2: Behavioral results a. Participants were at near-ceiling performance throughout the main task, $\mu = 0.905, \sigma = 0.05$. **b.** Participants reacted faster the higher the EV (x-axis) and slower to incongruent (purple) compared to congruent (green) trials. An interaction of EV $\times$ Congruency indicated stronger Congruency effect for higher EV ($p = .037$). Error bars represent corrected within subject SEMs [39, 40]. **c.** The Congruency effect was modulated by $\mathrm{EV}_{\mathrm{back}}$, i.e. the more participants could expect to receive from the ignored context, the slower they were when the contexts disagreed and respectively faster when contexts agreed (x axis, shades of colours). Error bars represent corrected within subject SEMs [39, 40]. **d.** Hierarchical model comparison for the main sample showed that including Congruency ($p < .001$), yet not $\mathrm{EV}_{\mathrm{back}}$ ($p = .27$), improved model fit. Including then an additional interaction of Congruency $\times$ $\mathrm{EV}_{\mathrm{back}}$ improved the fit even more ($p < .001$). **e.** We replicated the behavioral results in an independent sample of 21 participants outside of the MRI scanner. Including Congruency ($p = .009$), yet not $\mathrm{EV}_{\mathrm{back}}$ ($p = .63$), improved model fit. Including an additional interaction of Congruency $\times$ $\mathrm{EV}_{\mathrm{back}}$ explained the data best ($p = .017$).

A baseline model including only the factor EV indicated that participants reacted faster in trials that yielded bigger rewards ($\chi^2_{(1)} = 1374.6$, $p < .001$, Fig. 2b), in line with previous literature [41–43]. In the first step, we added either Congruency or $\mathrm{EV}_{\mathrm{back}}$ to the model. We found that Congruency also affected RTs, i.e. participants reacted slower to incongruent compared to congruent trials (t-test: $t_{(39)} = 4.59$, $p < .001$, likelihood ratio test to asses improved model fit: $\chi^2_{(1)} = 29.9$, $p < .001$, Fig. 2b). Interestingly, neither adding a main effect for $\mathrm{EV}_{\mathrm{back}}$ nor the interaction of EV $\times$ $\mathrm{EV}_{\mathrm{back}}$ improved model fit (LR-test with added terms: $\chi^2_{(1)} = 1.21$, $p = .27$ and $\chi^2_{(1)} = .01$, $p = 0.9$ respectively), meaning neither larger irrelevant values, nor their similarity to the objective value influenced participants' behavior.

In a second step, we investigated if the Congruency effect represents merely an agreement between the contexts, or if it interacted with the expected value of the best choice in the other context, i.e the points associated with the most valuable irrelevant stimulus feature ($\mathrm{EV}_{\mathrm{back}}$). Indeed, we found that the higher $\mathrm{EV}_{\mathrm{back}}$ was, the faster participants were on congruent trials. In incongruent trials, however, higher $\mathrm{EV}_{\mathrm{back}}$ had the opposite effect (Fig. 2c, LR-test of model with added interaction: $\chi^2_{(1)} = 18.19$, $p < .001$). We found no effect of the value associated with the other, lower valued irrelevant feature that would not have been chosen (LR-test to baseline model: $\chi^2_{(1)} = 0.92$, $p = .336$), nor did it interact with Congruency ($\chi^2_{(1)} = 2.76$, $p = .251$). This means that the expected value of a 'counterfactual' choice resulting from consideration of the irrelevant features mattered, i.e. that the outcome such a choice could have led to, also influenced reaction times. The hierarchical model comparison is summarized in Fig. 2d. All the effects above also hold when running the models nested across the levels of EV (as well as Block and Context, see Fig. S2). All nuisance regressors had a significant effect on RT (all ps< 0.03 in the baseline model).

The main behavioral results were replicated in an additional sample of 21 participants that were tested outside of the MRI scanner (LR-tests: Congruency, $\chi^2_{(1)} = 6.89$, $p = .009$, ,$\mathrm{EV}_{\mathrm{back}}$,$\chi^2_{(1)} = .23$, $p = .63$, Congruency $\times$ $\mathrm{EV}_{\mathrm{back}}$, $\chi^2_{(1)} = 5.69$, $p = .017$,Fig.2e).

We note that similar to the $\mathrm{EV}_{\mathrm{back}}$ $\times$ Congruency interaction, we also found that higher EV slightly increased the Congruency effect (Fig. 2b, LR-test: $\chi^2_{(1)} = 4.34$, $p = .037$). However, the interaction of Congruency $\times$ EV did not survive model comparison in the replication sample ($\chi^2_{(1)} = 0.23$, $p = .63$).Alternative regression models considering for instance within-cloud or between-context value differences did not provide a better fit the RTs

(Fig.S3). An exploratory analysis investigating all possible 2-way interactions with all nuisance regressors can be found in Fig. S4.

We took a similar hierarchical approach to model accuracy of participants in 2D trials, using mixed effects models with the same nuisance regressors as in the RT analysis. This revealed a main effect of EV (baseline model: $\chi^2_{(1)} = 14.71$, $p < .001$), indicating higher accuracy for higher EV. Introducing Congruency and then an interaction of Congruency $\times$ EV_{back} further improved model fit (LR-test: $\chi^2_{(1)} = 66.12$, $p < .001$, $\chi^2_{(1)} = 6.99$, $p = .03$, respectively), reflecting decreased performance on Incongruent trials, with higher error rates occurring on trials with higher EV_{back}. Unlike RT, error rates were not modulated by the interaction of EV and Congruency (LR-test with EV $\times$ Congruency: $\chi^2_{(1)} = 0.05$, $p = .825$). Out of all nuisance regressors, only switch had an influence on accuracy ($\chi^2_{(1)} = 10.22$, $p = .001$, in the baseline model) indicating increasing accuracy with increasing trials since the last switch trial.

In summary, these results indicated that participants did not merely perform a value-based choice among features on the currently relevant context. Rather, both reaction times and accuracy indicated that participants also retrieved the values of irrelevant features and computed the resulting counterfactual choice.

fMRI results

Decoding multivariate value signal from vmPFC Our MRI analyses focused on understanding the impact of irrelevant reward expectations on value signals in vmPFC. We therefore first sought to identify a value-sensitive region of interest (ROI) that reflected expected values in 1D and 2D trials, following common procedures in the literature [e.g. 4]. Specifically, we analyzed the fMRI data using general linear models (GLMs) with separate onsets and EV parametric modulators for 1D and 2D trials (at stimulus presentation, see online methods for full model). The union of the EV modulators for 1D and 2D trials defined a functional ROI for value representations that encompassed 998 voxels, centered on the vmPFC (Fig. 3a, $p < .0005$, smoothing: 4mm, to match the multivariate analysis), which was transformed to individual subject space for further analyses (mean number of voxels: 768.14, see online methods).

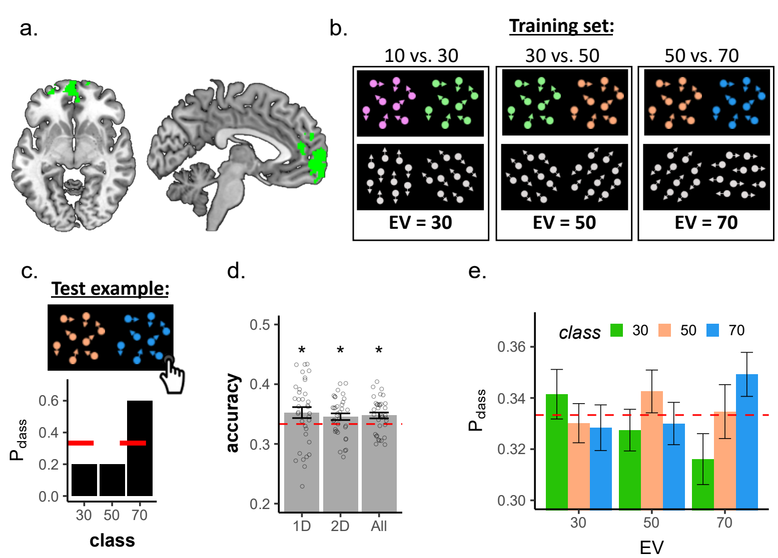


Figure 3: Multivariate value analyses. a. The union of the EV parametric modulator allowed us to isolate a cluster in the vmPFC. Displayed coordinates in the figure: x=-6, z=-6. **b.** We trained the classifier on behaviorally accurate 1D trials on patterns within the functionally-defined vmPFC ROI. **c.** The classifier yielded for each testing example one probability for each class. **d.** The classifier assigned the highest probability to the correct class (objective EV) significantly above chance for 1D trials, but also generalized to 2D and across all trials ($p = .049$, $p = .039$, $p = .007$ respectively). Error bars represent corrected within subject SEMs [39, 40]. **e.** Analyses of all probabilities revealed gradual value similarities. The y-axis represents the probability assigned to each class, colors indicate the classifier class and the x-axis represents the trial type (the objective EV of the trial). As can be seen, the highest probability was assigned to the class corresponding to the objective EV of the trial. Error bars represent corrected within subject SEMs [39, 40]

In the next step we focused on the multivariate activation patterns in the above-defined functional ROI. We trained a multivariate multinomial logistic regression classifier to distinguish the EVs of accurate 1D trials based on fMRI data acquired approximately 5 seconds after stimulus onset (Fig. 3b; leave-one-run-out training; see online methods for details). For each testing example, the classifier assigned the probability of each class given the data (i.e. '30','50' and '70', which sum up to 1, Fig. 3c). Because the ROI was constructed such as to contain significant information about EVs, the classifier should predict the correct EV. As expected, the class with the maximum probability corresponded to the objective outcome more often than chance in 1D trials ($\mu_{1D} = .35, \sigma_{1D} = .054$). Importantly, EV decoding also generalized to a test set composed of 1D and 2D trials ($\mu_{all} = .35, \sigma_{all} = .029$, $t_{(34)} = 2.89$, $p = .007$), and was significant when testing only on 2D trials ($\mu_{2D} = .35, \sigma_{2D} = .033$, $t_{(34)} = 2.20$,$p = .034$, Fig. 3d), even though the training data was restricted to 1D trials.

The following analyses model directly the class probabilities estimated by the classifier. Probabilities were modelled with beta regression mixed effects models [44]. For technical reasons, we averaged across nuisance regressors used in behavioral analyses. An exploratory analysis of raw data including nuisance variables showed that they had no influence and confirmed all model comparison results reported below (see Fig S6 and S8).

Multivariate neural value codes reflect value similarities and are negatively affected by contextually-irrelevant value information. We next asked whether EVs affected not only the probability of the corresponding class, but also influenced the full probability distribution predicted by the classifier. We reasoned that if the classifier is decoding the neural code of values, then similarity between the values assigned to the classes will yield similarity in probabilities associated to those classes. Specifically, we expected not only that the probability associated with the correct class be highest (e.g. '70'), but also that the probability associated with the closest class (e.g. '50') would be higher than the probability with the least similar class (e.g. '30', Fig. 3e). To test our hypothesis, we modelled the probabilities in each trial as a function of the absolute difference between the objective EV of the trial and the class (|EV-class|, i.e. in the above example with a correct class of 70, the probability for the class 50 will be modelled as condition 70-50=20 and the probability of 30 as 70-30=40). This analysis indeed revealed such a value similarity effect ($\chi^2_{(1)} = 12.74$, $p < .001$) also when tested separately on 1D and 2D trials ($\chi^2_{(1)} = 14.22$, $p < .001$, $\chi^2_{(1)} = 9.99$, $p = .002$, respectively, Fig. 4a). We compared this value similarity model to a perceptual model that merely encodes the amount of perceptual overlap between each training class and 2D testing (irrespective of their corresponding values) and found that our model explained the data best (Fig. 4b and Fig. S6).

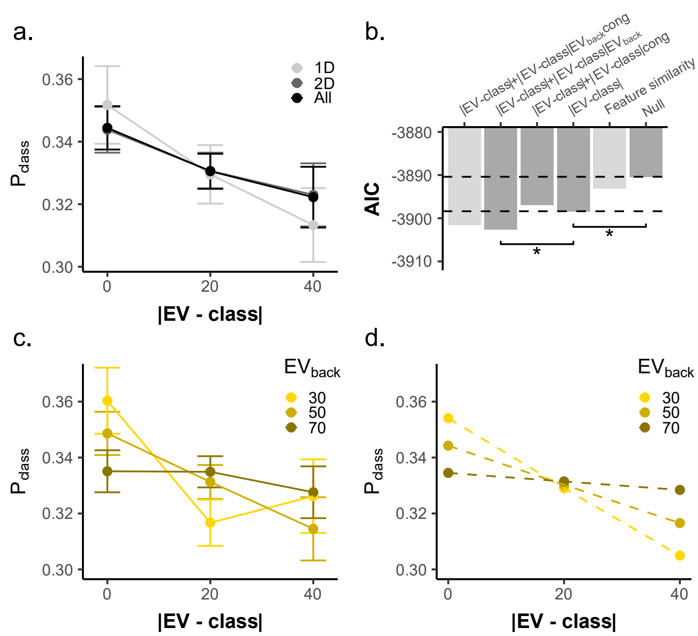


Figure 4: Impact of irrelevant feature values on value representations in vmPFC. a. Larger difference between the decoded class and the objective EV of the trial (x axis) was related to a lower probability assigned to that class (y axis) when tested in 1D, 2D or all trials (all $p < .002$, grey shades). Hence, the multivariate classifier reflected gradual value similarities. Note that when |EV - class|=0, P_{class} is the probability assigned to the objective EV of the trial. Error bars represent corrected within subject SEMs [39, 40] **b.** AIC values of competing models of value probabilities classified from vmPFC. Hierarchical model comparison of 2D trials revealed not only the differences between decoded class and objective EV (|EV-class|) improved model fit ($p < .002$), but rather that EV_{back} modulated this effect ($p = .013$). Crucially, Congruency did not directly modulate the value similarity ($p = .446$). Light gray bars represent models outside the hierarchical comparison. Including a 3-way interaction (with both EV_{back} and Congruency) did not provide better AIC score. A perceptual model encoding the feature similarity between each testing trial and the training classes (irrespective of values) did not provide a better AIC score than the value similarity model (|EV-class|). **c-d**. The higher the EV_{back} was, the weaker the effect of value similarity on the classifier's probabilities ($p = .013$). Data presented in (c) and model in (d). Error bars represent corrected within subject SEMs [39, 40].

Our main hypothesis was that context-irrelevant values might influence neural codes of value in the vmPFC. The experimentally manipulated background values in our task should therefore interact with the EV probabilities decoded from vmPFC. We thus tested the EV classifier only on 2D trials and asked whether the above described value similarity effect was influenced by EV$_{\text{back}}$ and\or Congruency. Analogous to our RT analyses, we used a hierarchical model comparison approach and tested if the interaction of value similarity with these factors improved model fit, using χ^2 based LR-tests (Fig. 4b). We found that EV$_{\text{back}}$, but not Congruency, modulated the value similarity effect ($\chi^2_{(1)} = 6.16$, $p = .013$, $\chi^2_{(1)} = .58$, $p = .446$, respectively, Fig. 4c). This effect indicated that the higher the EV$_{\text{back}}$ was, the less steep was the value similarity effect. Although including a 3-way interaction also improved model fit over a baseline model (Congruency $\times$ EV$_{\text{back}}$ $\times$ |EV-class|, $\chi^2_{(1)} = 7.2$,$p = .027$), the AIC score did not surpass the model with only the 2-way interaction (-3902.5,-3901.6, respectively). These results also hold when running the models nested within the levels of EV (Fig.S6). Replacing the EV$_{\text{back}}$ with a parameter that encodes the presence of the perceptual feature corresponding to EV$_{\text{back}}$ in the training class (Similarity$_{\text{back}}$: 1 if the feature was preset, 0 otherwise, see Fig. S7) did not provide a better AIC score (-3897.1) than including the value of EV$_{\text{back}}$ (-3902.5). Note that main effects of EV$_{\text{back}}$ or Congruency would not be sensible to test in this analysis because both factors don't discriminate between the classes, but rather assign the same value to all three probabilities from that trial (which sum to 1).

In summary, this indicates that the neural code of value in the vmPFC is affected by contextually-irrelevant value expectations, such that larger alternative values disturb neural value codes in vmPFC more than smaller ones. This was the case even though the alternative value expectations were not relevant in the context of the considered trials. The effect occurred irrespective of the agreement or action-conflict between the relevant and irrelevant values, unlike participants' behaviour, which were mainly driven by Congruency and it's interaction with EV$_{\text{back}}$. Our finding suggests that the (counterfactual) value of irrelevant features must have been computed and poses the power to influence neural codes of objective EV in vmPFC.

Larger irrelevant value expectations are related to reduced relevant EV signals, influencing behavior. While modelling the full probability distribution over values offers important insights, it only indirectly sheds light on the neural representation of the objective EV that reflects participants' choices in correct trials. We next focused on modelling the probability associated with the class corresponding to the objective EV of each 2D trial (henceforth: P_{EV}). This also resolved the statistical issues arising from the dependency of the three classes (i.e. for each trial they sum to 1). As can be inferred by Fig 3e above, the median probability of the objective EV on 2D trials was higher than the the average of the other non-EV probabilities ($t_{(34)} = 2.50$, $p = .017$) In line with the findings reported above, we found that EV$_{\text{back}}$ had a negative effect on P_{EV} ($\chi^2_{(1)} = 5.96$, $p = .015$, Fig. 5a), meaning that higher EV$_{\text{back}}$ was associated with a lower probability of the objective EV, P_{EV}. Interestingly, and unlike in the behavioral models, we found that neither Congruency nor its interaction with EV or with EV$_{\text{back}}$ influenced P_{EV} ($\chi^2_{(1)} = 0.035$, $p = .852$,$\chi^2_{(1)} = 0.48$, $p = .787$, $\chi^2_{(1)} = .99$, $p = .317$, respectively, Fig. 5b). The effect of EV$_{\text{back}}$ also holds when running the model nested inside the levels of EV ($\chi^2_{(1)} = 5.99$, $p = 0.014$, Fig.S8b). A model including an additional regressor that encoded trials in which EV=EV$_{\text{back}}$ (or: match) did not improve model fit, and no evidence for an interaction of the match regressor with the EV$_{\text{back}}$ was found (LR test with added terms: $\chi^2_{(1)} = 0.45$, $p = .502$, $\chi^2_{(1)} = 0.77$, $p = .379$, respectively). This might indicate that when value expectations of both contexts matched, there was neither an increase nor a decrease of P_{EV}. Lastly, we verified that replacing EV$_{\text{back}}$ with the perception-based Similarity$_{\text{back}}$ regressor did not provide a better model fit (AICs: -1229.2,-1223.3, respectively). These findings confirm that EV$_{\text{back}}$ is not only disturbing the neural code of values in the vmPFC but also specifically decreases the decodability of the objective EV.

As in our behavioral analysis, we evaluated alternative models of P_{EV} that included a factor reflecting within-option or between-context value differences, or alternatives for EV$_{\text{back}}$ (Fig.S8). This exploratory analysis revealed that our model provides the best fit for P_{EV} in all cases except when EV$_{\text{back}}$ was replaced with the sum of irrelevant values (-1229.6, -1229.2, respectively, Fig. S8). In contrast, AIC scores of behavioral models' favored EV$_{\text{back}}$ as modulator of Congruency, over the sum of irrelevant values (-6626.6, -6619.9, respectively, Fig.S3). However, both

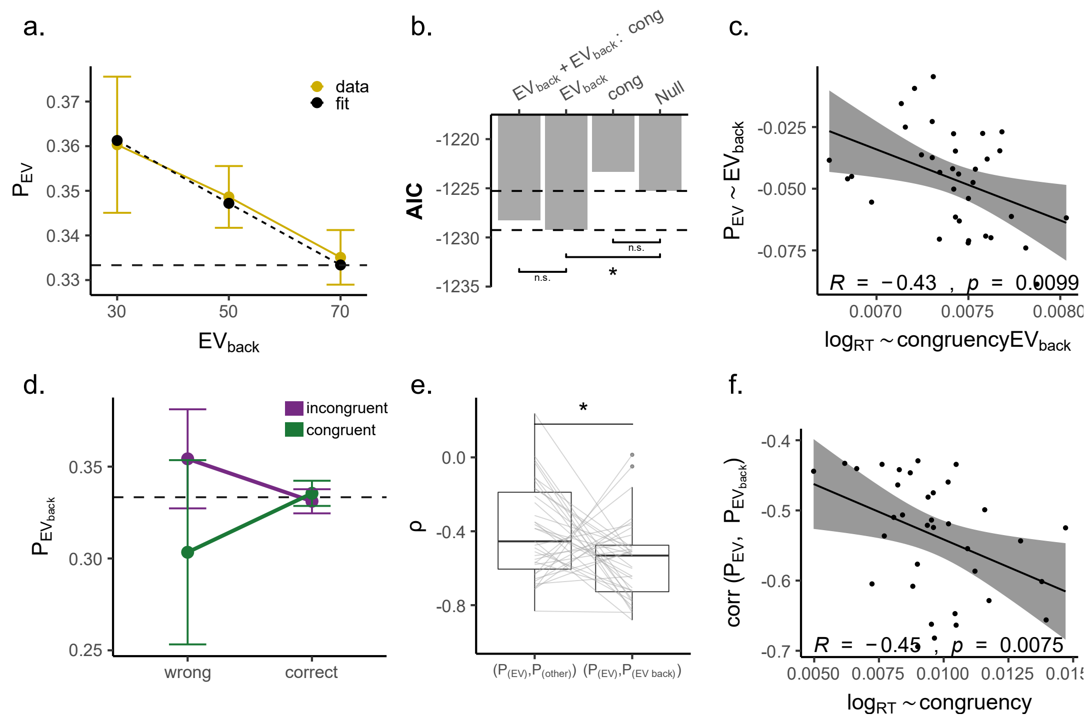


Figure 5: Multivariate results: decoding the EV a. Higher EV_{back} was related to a decreased decodability of EV ($p = .015$). Yellow line reflects data, dashed line model fit from mixed effects models described in text. Error bars represent corrected within subject SEMs [39, 40]. **b.** Hierarchical model comparisons revealed that the effect of EV_{back} alone explained data best ($p = .015$) and no main effect or interaction with Congruency was indicated (Congruency main effect, $p = .852$, Congruency $\times$ EV_{back}, $p = .317$). **c.** Participants who had a stronger effect of EV_{back} on the EV decodability (y-axis, more negative values indicate stronger decrease of P_{EV} as a result of EV_{back}, see panel a) also had a stronger modulation of EV_{back} on the effect of Congruency on their RT (x-axis, more positive values indicate stronger influence on the slow incongruent and fast congruent trials). **d.** The probability associated with EV_{back} ($\text{P}_{\text{EV}_{\text{back}}}$, y-axis) was increased when participants chose the option based on EV_{back}. Specifically, in incongruent trials (purple), high $\text{P}_{\text{EV}_{\text{back}}}$ was associated a wrong choice, whereas in Congruent trials (green) it was associated with correct choices. This effect is preserved when modeling only wrong trials (main effect of Congruency: $\chi^2_{(1)} = 4.36$, $p = .037$). Error bars represent corrected within subject SEMs [39, 40]. **e.** The correlation of P_{EV} and $\text{P}_{\text{EV}_{\text{back}}}$ was stronger than with P_{Other}, $p = .017$. **f.** Participant that had a stronger (negative) correlation of P_{EV} and $\text{P}_{\text{EV}_{\text{back}}}$ (x-axis, more negative values indicate stronger negative relationship) also had a stronger effect of Congruency on their RT (y-axis, larger values indicate a stronger RT decrease in incongruent compared to congruent trials)

parameters were strongly correlated ($\rho = .87$, $\sigma = .004$) and therefore our task was not designed to distinguish between these two alternatives.

If the effect of EV_{back} indeed reflects an influence of contextually-irrelevant values on neural representations of the relevant expected value, then this might impact participants' behavior. We therefore asked whether the influence on the representation in vmPFC might relate to participants' reaction times. In line with this idea, we found that participants with a stronger EV_{back} effect on P_{EV} also had a stronger EV_{back} $\times$ Congruency interaction effect on their RT ($r = -.43$,$p = .01$, Fig. 5c).

Next, we tested whether vmPFC represents EV_{back} directly. A classifier trained on accurate 2D trials with the labels of EV_{back} could not successfully detect the correct class (t-test against chance: $t_{(34)} = 0.73$, $p = .47$). Note, however, that 2D trials were not fully balanced across the values of EV_{back} (Fig. 1e), which complicated obtaining enough trials for classifier training. We thus turned to look at the probability the classifier trained on 1D trials assigned to the class corresponding to EV_{back} (henceforth: $\text{P}_{\text{EV}_{\text{back}}}$). When focusing only on behaviorally accurate trials, we found no effect of EV nor Congruency on $\text{P}_{\text{EV}_{\text{back}}}$ ($\chi^2_{(1)} = 0.07$, $p = .794$, $\chi^2_{(1)} = 0.00$, $p = .987$ respectively). However, motivated by our behavioral analyses that indicated an influence of the irrelevant context on accuracy, we asked whether $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ was different on behaviorally wrong or incongruent trials. We found an interaction of accuracy × Congruency ($\chi^2_{(1)} = 4.51$, $p = .034$, Fig. 5d) that indicated increased $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ for accurate congruent trials and a decrease for wrong incongruent trials. Effectively, this means that in trials in which participants erroneously chose the option with higher valued irrelevant features, $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ was increased.

Parallel representation of task-relevant and task-irrelevant expected values in vmPFC. Our previous analyses indicated that the probability of the objective EV decreased with increasing $\mathrm{EV}_{\mathrm{back}}$. This decrease could reflect a general disturbance of the value retrieval process caused by the distraction of competing values. Alternatively, if the irrelevant values are represented within the same neural code as the objective EV, then the probability assigned to the class corresponding to $\mathrm{EV}_{\mathrm{back}}$ would increase in exchange for a decrease in P_{EV} – even though the classifier was trained in the absence of task-irrelevant values, i.e. the objective EV of 1D trials. In order to test this idea, we took the same trained classifier and tested it only on trials in which $\mathrm{EV} \neq \mathrm{EV}_{\mathrm{back}}$, i.e. in which the value expected in the current task context was different than the value that would be expected for the same choice in a different task-context. This allowed us to re-label the classes of each trial to P_{EV}, $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ and $\mathrm{P}_{\mathrm{other}}$, where 'other' corresponds to the class that is neither the EV nor $\mathrm{EV}_{\mathrm{back}}$ of the trial, and examine directly the correlation between each pair of classes. To prevent a bias between the classes, we only included trials in which the class corresponding to 'other' appeared on the screen as either relevant or irrelevant value.

For each trial, the three class probabilities sum up to 1 and hence are strongly biased to correlate negatively with each other. Not surprisingly, we found such strong negative correlations across participants of both pairs of probabilities, i.e. between P_{EV} and $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ ($\rho = -.56$, $\sigma = .22$) as well as between P_{EV} and $\mathrm{P}_{\mathrm{other}}$ ($\rho = -.40$, $\sigma = .25$). However, we found that the former correlation was significantly stronger than the latter ($t_{(34)} = -2.77$, $p = .017$, Fig. 5e), indicating that when the probability assigned to the EV decreased, it was accompanied by a stronger increase in the probability assigned to $\mathrm{EV}_{\mathrm{back}}$, akin to a competition between both types of expectations. Additionally, a formal model predicting P_{EV} by $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ resulted in a smaller (i.e. better) AIC (-567.13), compared to using $\mathrm{P}_{\mathrm{other}}$ as predictor (-475.32, see online methods). In line with this finding, we turned to test if this potential competition is reflected in participants' behavior. Of particular relevance in this regard is the behavioral Congruency effect, which similarly reflects a competition between the different values. Strikingly, we found that the more negatively P_{EV} correlated with $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$, the stronger Congruency influenced participants' behavior ($r = -.45$,$p = .008$, Fig. 5f).

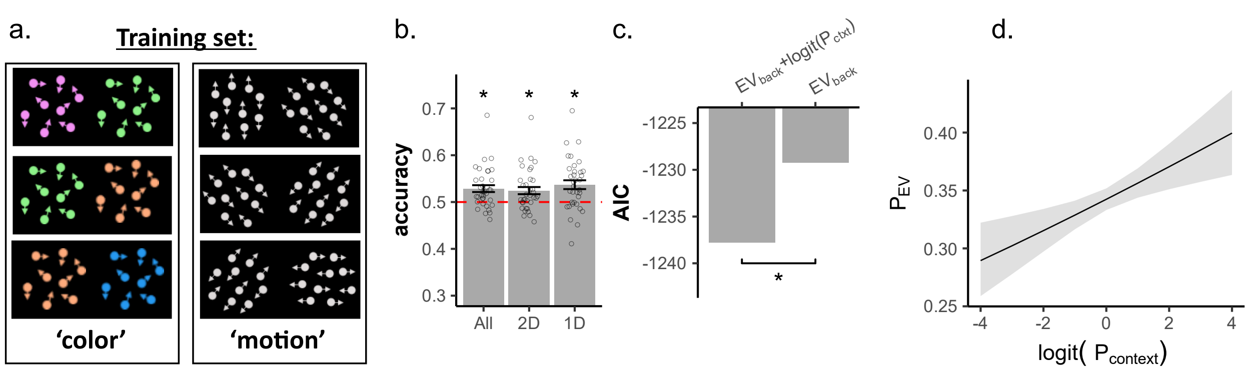


Figure 6: Context decodability in the vmPFC directly relates to the representation of the objective outcome a. We trained the same classifier on the same data only this time we split the training set to classes corresponding to the two possible contexts: Color (left) or Motion (right), irrespective of the EV, though we kept the training sets balanced for EV (see online methods). **b.** The classifier could decode the trial's context above chance also when sub-setting the data to 1D, 2D and when testing on all trials ($p < .001$,$p = .002$,$p < .001$, respectively). Error bars represent corrected within subject SEMs [39, 40] **c.** The trial-context decodability improved prediction of the objective outcome probability, beyond the $\mathrm{EV}_{\mathrm{back}}$ ($p = .001$). **d.** The objective outcome was strongly represented (P_{EV}), the more the context was decodable from the vmPFC (modeled as logit-transformed probability assigned to the trial-context of the trial, x-axis)

In summary, the neural code in vmPFC is mainly influenced by the contextually relevant EV. However, if an alternative context would lead to a large expected value, the representation of the relevant expected value is weakened, irrespective of their agreement on the action to be made. Moreover, weakening of the EV representation is accompanied by a strengthening of the representation of $\mathrm{EV}_{\mathrm{back}}$ on a trial by trial basis. Lastly, participants with a stronger influence of high alternative values on the EV representation also had a stronger influence of $\mathrm{EV}_{\mathrm{back}}$ on the Congruency RT effect. Likewise, participants who exhibited a larger negative association between the decodability of EV and the decoded probability of $\mathrm{EV}_{\mathrm{back}}$, also reacted slower when the contexts pointed to different actions. As will be discussed later in detail, we consider this to be evidence for parallel processing of two task aspects in this region, EV and $\mathrm{EV}_{\mathrm{back}}$.

Task-context representations interact with value codes within vmPFC Above we reported that vmPFC activity is influenced by multiple value expectations. Which value expectation is currently relevant depended on the task context. We therefore hypothesized that, in line with previous work, vmPFC would also encode the task context, although this is not directly value-related. We thus turned to see if we can decode the trial's context from the same region that was univariately sensitive to EV. For this analysis we trained the same classifier on the same accurate 1D trials as before, only it was trained to distinguish the trial types 'Color' and 'Motion' (Fig. 6a). Crucially, the classifier had no information as to what was the EV of each given trial, and training sets were up-sampled to balance the EVs within each set (see online methods). The classifier was above chance for decoding the correct context in 1D, 2D and all trials ($t_{(34)} = 3.95$, $p < .001$, $t_{(34)} = 3.2$, $p = .003$, $t_{(34)} = 3.93$, $p < .001$, respectively, Fig.6b). Additionally, the context is decodable also when only testing on 2D trials in which value difference in both contexts was the same (i.e. when keeping the value difference of the background 20, since the value difference of the relevant context was always 20, $t_{(34)} = 2.73$, $p = .01$).

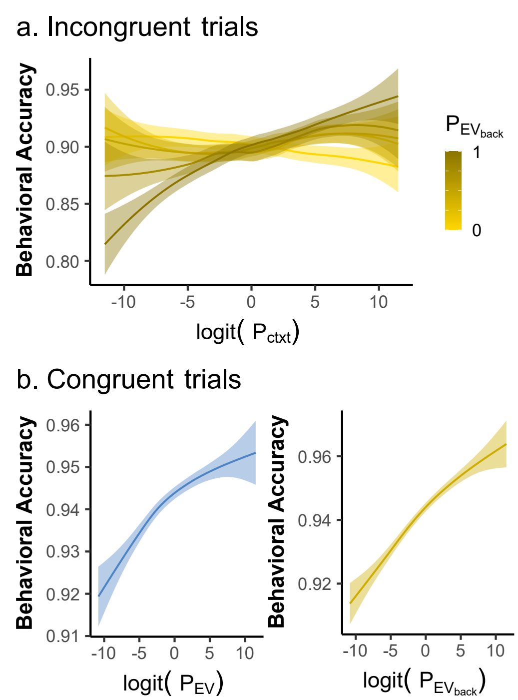


Figure 7: Neural representations of context and value in vmPFC jointly guide behavior a. Lower context decodability of the relevant context (x axis) was associated with less behavioral accuracy (y-axis) in incongruent trials ($p = .051$). This effect was modulated by the representation of $\mathrm{EV}_{\mathrm{back}}$ in vmPFC ($p = .012$, shades of gold), i.e. it was stronger in trials where $\mathrm{EV}_{\mathrm{back}}$ was strongly decoded from the vmPFC (shades of gold, plotted in 5 quantiles). Shown are fitted slopes from analysis models reported in the text. **b.** Decodability of both EV ($p = .058$, blue, left) and $\mathrm{EV}_{\mathrm{back}}$ ($p = .009$, gold, right) had a positive relation to behavioral accuracy (y axis) in congruent trials. Shown are fitted slopes from analysis models reported in the text.

Importantly, if vmPFC is involved in signaling the trial context as well as the values, then the strength of context signal might relate to the strength of the contextually relevant value. Strikingly, we found that $\mathrm{P}_{context}$ had a positive effect on the decodability of EV and that adding this term in addition to $\mathrm{EV}_{\mathrm{back}}$ to the P_{EV} model improved model fit ($\chi^2_{(1)} = 10.5$, $p = .001$, Fig. 6c-d). In other words, the more the context was decodable, the higher was the probability assigned to the correct EV class.

Lastly, we investigated how neural representations in vmPFC of EV, $\mathrm{EV}_{\mathrm{back}}$ and the relevant Context influence participants' accuracy. Note that the two contexts only indicate different choices in incongruent trials, where a wrong choice might be a result of a strong influence of the irrelevant context. The behavioral effect on accuracy could therefore be particularly relevant in this condition. This was also indicated by the analysis of $\mathrm{P}_{\mathrm{EV}_{\mathrm{back}}}$ shown in Fig 5d. We therefore modeled congruent and incongruent trials separately. This showed that that a weaker representation of the relevant context was marginally associated with an increased error rate (negative effect of P$_{context}$) on accuracy, LR-test with P$_{context}$): $\chi^2_{(1)} = 3.66$, $p = .055$). Moreover, if stronger representation of the wrong context (i.e. 1-P$_{context}$)) is reducing accuracy, than stronger representation of the value associated with this context (EV$_{\mathrm{back}}$) should strengthen that influence. Indeed, we found that adding a P$_{context}$ $\times$ P$_{\mathrm{EV_{back}}}$ term to the model explaining error rates improved model fit ($\chi^2_{(1)} = 6.33$, $p = .012$, Fig. 7a). Yet, the representation of EV and EV$_{\mathrm{back}}$ did not directly influence behavioral accuracy (P$_{\mathrm{EV}}$: $\chi^2_{(1)} = 0.28$, $p = .599$,P$_{\mathrm{EV_{back}}}$: $\chi^2_{(1)} = 0.0$, $p = .957$). In congruent trials choosing the wrong choice is unlikely a result of wrong context encoding, since both contexts lead to the same choice. Indeed, there was no influence of P$_{context}$) on accuracy for congruent trials (LR-test: $\chi^2_{(1)} = 0.0$, $p = .922$). However, strong representation of either relevant or irrelevant EV would lead to a correct choice.Indeed, we found that both an increase in P$_{\mathrm{EV_{back}}}$ and (marginally) in P$_{\mathrm{EV}}$ had a positive relation to behavioral accuracy (P$_{\mathrm{EV_{back}}}$: $\chi^2_{(1)} = 6.48$, $p = .011$, P$_{\mathrm{EV}}$: $\chi^2_{(1)} = 3.5$, $p = .061$, Fig. 7b).

No evidence for univariate modulation of contextually irrelevant information on expected value signals in vmPFC The above analyses indicated that multiple value expectations are represented in parallel within vmPFC. Lastly, we asked whether whole-brain univariate analyses could also uncover evidence for processing of multiple value representations. In particular, we asked whether we could find evidence for a single representation that integrates the multiple value expectations into one signal. To this end, we first analyzed the fMRI data using GLMs with separate onsets and EV parametric modulators for 1D and 2D trials (see online methods for details). As expected, several regions were modulated by EV in both trial types, including vmPFC (EV$_{\mathrm{1D}} > 0 \cap$ EV$_{\mathrm{2D}}$ >0, Fig.8a). Hence, the vmPFC signaled the expected value of the current context in both trial types as expected – even though 2D trials likely required higher attentional demands (and indeed, the attention network was identified for the 2D>1D contrast, p<.001, Fig.8b)

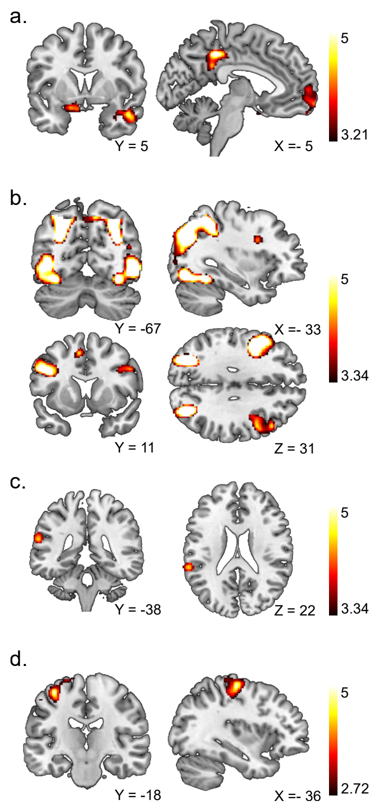


Figure 8: Univariate results Depicted are T-maps for each contrast. A detailed table of clusters can be found in the SI S1. **a**. The intersection of the EV parametric modulators of 1D and 2D trials revealed several regions including right Amygdala, bilateral Hippocampus and Angular Gyrus, the lateral and medial OFC and overlapping vmPFC. Voxelwise threshold $p < .001$, FDR cluster-corrected. **b** 2D trials were characterized by increased activation in an attentional network involving occipital, parietal and frontal clusters (2D $>$ 1D, $p < .001$ FDR cluster corrected). **c.** A region in the Superior Temporal Gyrus was negatively modulated by EV$_{\mathrm{back}}$, i.e. the higher the EV$_{\mathrm{back}}$, the lower the signal in this region. $p < .001$, FDR cluster-corrected. No overlap with (b), see S9. **d.** A cluster in the primary motor cortex was negatively modulated by Congruency $\times$ EV$_{\mathrm{back}}$, i.e. the difference between Incongruent and Congruent trials increased with higher EV$_{\mathrm{back}}$, similar to the RT effect, $p < .005$, FDR cluster-corrected. No overlap with (b), see S9

Next, we searched for univariate evidence for processing of irrelevant values by modifying the parametric modulators assigned to 2D trials in the above-mentioned GLM (for full models, see Fig S9). Specifically, in addition to EV$_{\mathrm{2D}}$, we added Congruency (+1 for congruent and -1 for incongruent) and EV$_{\mathrm{back}}$ as additional modulators of the activity in 2D trials. This GLM revealed no evidence for a Congruency contrast anywhere in the brain (even at a liberal voxel-wise threshold of $p < .005$), but an unexpected negative effect of EV$_{\mathrm{back}}$ in the Superior Temporal Gyrus ($p < .001$, Fig.8c). Notably, unlike the multivariate analysis, no effect in any frontal region was observed. Motivated by our behavioral analysis, we then turned to look for the interaction of each relevant or irrelevant value with Congruency. An analysis including only a Congruency $\times$ EV$_{\mathrm{2D}}$ parametric modulator revealed no cluster (even at $p < .005$). Another analysis including Congruency $\times$ EV$_{\mathrm{back}}$ in addition to EV$_{\mathrm{2D}}$ as parametric modulators, however, revealed a negative effect in the primary motor cortex at a liberal threshold, which indicated that the difference between Incongruent and Congruent trials increased with higher EV_{back}, akin to a response conflict ($p < .005$, Fig.8d). Lastly, we re-ran all above analyses concerning Congruency and EV_{back} only inside the identified vmPFC ROI. No voxel survived for Congruency, EV_{back} nor the interactions, even at threshold of $p < .005$.

Additional exploratory analyses such as contrasting the onsets of congruent and incongruent trials, confirmed the lack of Congruency modulation in any frontal region (Fig. S9). Interestingly, at a liberal threshold of $p < .005$ we found stronger activity for 1D over 2D trials in a cluster overlapping with vmPFC (1D > 2D, $p < .005$, S9). Although this could be interpreted as a general preference for 1D trials, splitting the 2D onsets by Congruency revealed no cluster for 1D > Incongruent (also at $p < .005$) but a stronger cluster for 1D > Congruent ($p < .001$,Fig. S9). In other words, the signal in the vmPFC was *weaker* when both contexts indicate the same action, compared to when only one context is present.

In summary, our univariate analyses indicated the well-known sensitivity of vmPFC to values expected within the relevant context. Yet, unlike our multivariate analyses, we found no evidence for signal modulation by contextually irrelevant values outside the motor cortex, where we found a negative modulation of Congruency $\times$ EV_{back}. This contrasts with the idea that competing values would have been integrated into a single EV representation in the vmPFC, because this account would have predicted a higher signal for Congruent compared to Incongruent trials. If at all, we found a general decrease in signal for Congruent trials.

Discussion

In this study, we investigated how contextually-irrelevant value expectations influence behavior as well as neural vmPFC activation patterns. We asked participants to make choices between options that had different expected values in different task-contexts. Participants reacted slower when the expected values in the irrelevant context favored a different choice, compared to trials in which relevant and irrelevant contexts favored the same choice. This Congruency effect increased with increasing reward associated with the hypothetical choice in the irrelevant context (EV_{back}). We then identified a functional ROI that is univariately sensitive to the objective expected values (EV), i.e. the contextually-relevant rewards. Multivariate analyses revealed that a high EV_{back} disrupts the value-code of the vmPFC. Specifically, higher EV_{back} was associated with a degraded representation of the objective EV (P_{EV}) in vmPFC. At the same time, increased representation of EV_{back} in the vmPFC during stimuli presentation was associated with an increased chance of choosing accordingly, irrespective of its agreement with the relevant context. Moreover, the decrease in decodability of the value in the relevant context was associated with an increase in the value that would be obtained in the other task-context ($P_{EV_{back}}$), akin to a conflict of the two value representations. Both these effects were associated with the congruency-related behavioral slowing. Importantly, we also found that the task context (color/motion) could be decoded from the same brain region. This decodability of the context was related to the decodability of the value in the current context. Lastly, we are aware that in binary decoding, low decodability of the correct class doesn't necessarily point to high decodability of the alternative class. Nevertheless, when the irrelevant context pointed to the wrong choice in incongruent trials, stronger vmPFC representation of the alternative (wrong) context and its corresponding value were related to higher error rates. However, when both contexts agreed on the action to be made, stronger representation of either of their EVs were strongly related to making a correct choice.

We found no evidence that the signal in vmPFC is sensitive to Congruency. The only region that was univariately modulated by Congruency was the primary motor cortex. These data suggest a complex multi-faceted value representation in vmPFC, in which multiple values of the same option under different task-contexts are reflected and influence behavior. While we could not directly decode EV_{back}, it had a significant and value-dependent effect on EV representations, hinting at a complex form of co-represention within the vmPFC. Moreover, we could also decode the current task-context from vmPFC, and the strength of context encoding is related to the strength of the representation of the context associated value.

Behavioral analyses showed outcome-irrelevant values are not completely filtered. In our experiment the relevant features were cued explicitly and the rewards were never influenced by the irrelevant features. Nevertheless, participants' reactions were influenced by not only the contextually relevant outcome, but also the counterfactual choice, based on values irrelevant in the given context. These results raise the question how internal value expectation(s) of the choice are shaped by the possible contexts. One hypothesis could be that rewards expected in both contexts integrate into a single EV for a choice, which in turn guides behavior. This perspective suggests that the expected value of options valuable in both contexts will increase, relative to options that are valuable only in the current but not in the alternative context. In other words, in trials in which the irrelevant context agreed with the decision, the (subjective) EV of choice might increase, in proportion to how large the irrelevant value was. However, if the alternative context disagrees, the (subjective) EV might decrease. This approach would treat RT as a direct measure of EV.

An alternative hypothesis would be that both values are kept separate, and will be processed in parallel. In this case, their conflict would have to be resolved in a different brain region, such as the motor cortex. This would suggest that behavior is guided by two value expectations that are resolved into action, likely outside the value-network. To differentiate these possibilities motivated us to focus our analysis on the vmPFC, where we could distinguish between a single integrated value and simultaneously oc-occuring representations. Notably, the interaction of values could be also influenced by a representation of the current task context, which is known to be represented in the same region and the overlapping orbitofrontal cortex [e.g., 32, 34, 35, 45]. It therefore seemed to be a good candidate region to help illuminate how values stemming from different contexts, as well as information about the contexts themselves, might interact in the brain.

The lack of a Congruency effect on univariate vmPFC signals contradicted the integration hypothesis. Even before considering the specific outcomes of the two contexts, it would predict an increased signal for congruent compared to incongruent trials. If at all, we find the univariate vmPFC activation in 1D trials to be stronger than in Congruent 2D trials.

Interestingly, the univariate analysis was not sensitive enough to detect an influence of the irrelevant values in vmPFC. Only an investigation into the multivariate analyses revealed a degraded EV representation in trials with stronger alternative values, suggesting that the two potential values are in representational conflict. This impact on value representations occurred irrespective of choice congruency, but correlated with the behavioral modulation of EV_{back} on congruency. Due to limitations of our design, we could not successfully train a classifier directly on EV_{back} of 2D trials. Moreover, the objective class was not strongly represented when both expected values (EV and EV_{back}) were the same, suggesting some differences in the underlying representations of relevant and irrelevant values. However, a classifier trained on EV in 1D trials in which no irrelevant values were present, was still sensitive to the expected value of the irrelevant context in 2D trials. This could suggest that within the vmPFC 'conventional' expected values and counterfactual values are encoded using at least partially similar patterns.

This interpretation would also be supported by our findings that both representations contributed to choice accuracy in Congruent trials, and that $P_{EV_{\mathrm{back}}}$ and P_{EV} were negatively correlated, such that decreases in the EV representation were accompanied by an increased EV_{back} representation. This might also explain how the reducing effect EV_{back} had on the EV representation aligns well with behavioral changes observed in incongruent trials (i.e. reducing both RT and accuracy), but also our finding of improved performance on congruent trials, even though there EV_{back} could still be large: in the first case, when choices for the two context differ, competing EV and EV_{back} lead to performance decrements; in the second case, when choices are the same, both of the independently contributing representations would support the same reaction and therefore benefit performance. Our results therefore are in line with the interpretation that both relevant and irrelevant values are retrieved, represented in parallel within the vmPFC and influence behavior.

At the same time, our results also suggest that while the $\mathrm{EV}_{\mathrm{back}}$ influenced the representation of EV, the latter largely dominated population activity. This is in line with our task requirements and participant's high behavioral accuracy that indicated accurate choices were driven by EV in the vast majority of cases. However, even when focusing only on behavioral accurate trials, we see that the signal in vmPFC encompass a representational conflict between the two EVs, which was related to Congruency-dependent RT effects in those trials. Interestingly, univariate analyses were not sensitive enough to detect an influence of the outcome-irrelevant values in the vmPFC.

Univariate analyses revealed a weak negative modulation of primary motor cortex activity by Congruency. Akin to a response conflict, this corresponds to recent findings that distracting information can be traced to areas involved in task execution cortex in humans and monkeys [21, 22]. Crucially however, unlike in previous studies the modulation found in our study was dependent on the specific values of the alternative context. This could suggest that the outcome-representation conflict in the vmPFC is resolved in the primary motor cortex. This would also be in line with our interpretation that the vmPFC does not integrate both tasks into a single EV representation.

One important implication of our study concerns the nature of neural representations in the vmPFC/mOFC. A pure perceptual representation should be equally influenced by all four features on the screen. Yet, our decoding results could not have been driven by the perceptual properties of the chosen feature, and effects of background values could also not be explained by perceptual features of the ignored context (Fig. 3 and Fig. S7). Moreover, we show that the signal in vmPFC reflects more than expected values of the choice, and we did not find any evidence for value integration. Finally, investigating trials on which both expected values, EV and $\mathrm{EV}_{\mathrm{back}}$, were the same, we did not find a stronger signal for the objective class. This indicates that our classifier was neither exclusively sensitive to the perceptual features, nor to values regardless of whether they were relevant or not. Both those accounts would predict an increased representation of the objective class in those trials. Instead, we show that vmPFC simultaneously represents option values as well as information about the current task-context, and that both these representations interact with each other as well as behavior. One possible solution which has been suggested in previous research is that vmPFC/mOFC might be tasked with representing a task-state, which effectively encodes the current state of all information relevant to the task, in particular if information is partially observable [32, 45]. Note that the task context, which we decode from vmPFC activity in the present paper, could be considered as a superset of the more fine grained task states that reflect the individual motion directions/colors involved in a comparison. Any area sensitive to these states would therefore also show decoding of context as defined here. Whether vmPFC has access to such detailed information about the states cannot be conclusively answered with the present research for power reasons.

Of note, some work has found that EV could be one additional aspect of OFC activity [36] that is multiplexed with other task-related information. Crucially, the idea of task-state as integration of task-relevant information [28, 46] could explain why this region was found crucial for integrating valued features, when all features of an object are relevant for choice [16, 28], although some work suggests that it might even reflect features not carrying any value [29]. Moreover, the link between context and EV decodability as well as to behavioral accuracy suggests a multi-faceted vmPFC representation which not only contains multiple values, but also links information about the relevant task context to the corresponding values, just as the task-state framework might suggest.

To conclude, the main contribution of our study is that we elucidated the relation between task-context and value representations within the vmPFC. By introducing multiple possible values of the same option in different contexts, we were able to reveal a complex representation of task structure in vmPFC, with both task-contexts and their associated values activated in parallel. The decodability of both context and value(s) independently from vmPFC, and their relation to choice behavior, hints at integrated computation of these in this region. We believe that this bridges between findings of EV representation in this region to the functional role of this region as representing task-states, whereby relevant and counterfactual values can be considered as part of a more encompassing state representation.

Acknowledgments

NWS was funded by an Independent Max Planck Research Group grant awarded by the Max Planck Society (M.TN.A.BILD0004) and a Starting Grant from the European Union (ERC-2019-StG REPLAY-852669). NM was funded by and is grateful for a scholarship from the Ernst Ludwig Ehrlich Studienwerk (ELES) and Einstein Center for Neuroscience (ECN) Berlin throughout this study. We thank Gregor Caregnato for help with participant recruitment, Anika Löwe, Lena Maria Krippner, Sonali Beckmann and Nadine Taube for help with data acquisition, all participants for their participation and the Neurocode lab for numerous contributions and help throughout this project.

Data availability statement

The MRI data that support the findings of this study will be made available upon publication.

Code availability statement

Custom code for all analyses conducted in this study will be made available upon publication.

Online Methods

Participants

Forty right-handed young adults took part in the experiment (18 women, $\mu_{age} = 27.6, \sigma_{age} = 3.35$) in exchange for monetary reimbursement. Participants were recruited using the participant database of Max-Planck-Institute for Human Development. Beyond common MRI-safety related exclusion criteria (e.g. piercings, pregnancy, large or circular tattoos etc.), we also did not admit participants to the study if they reported any history of neurological disorders, tendency for back pain, color perception deficiencies or if they had a head circumference larger than 58 cm (due to the limited size of the 32-channel head-coil). After data acquisition, we excluded five participants from the analysis; one for severe signal drop in the OFC, i.e. more than 15% less voxels in functional data compared to the OFC mask extracted from freesurfer parcellation of the T1 image [47, 48]. One participant was excluded due to excessive motion during fMRI scanning (more than 2mm in any axial direction) and three participants for low performance (less than 75% accuracy in one context in the main task). In the behavioral-replication, 23 young adults took part (15 women, $\mu_{age} = 27.1, \sigma_{age} = 4.91$) and two were excluded for the same accuracy threshold. Due to technical reasons, 3 trials (4 in the replication sample) were excluded since answers were recorded before stimulus was presented and 2 trials (non in the replication) in which RT was faster than 3 SD from the mean (likely premature response). The monetary reimbursement consisted of a base payment of 10 Euro per hour (8.5 for replication sample) plus a performance dependent bonus of 5 Euro on average. The study was approved the the ethics board of the Free University Berlin (Ref. Number: 218/2018).

Experimental procedures

Design Participants performed a random dot-motion paradigm in two phases, separated by a short break (minimum 15 minutes). In the first phase, psychophysical properties of four colors and four motion directions were first titrated using a *staircasing task*. Then, participants learned the rewards associated with each of these eight features during a *outcome learning task*. The second phase took place in the MRI scanner and consisted mainly of the *main task*, in which participants were asked to make decisions between two random dot kinematograms, each of which had one color and/or one direction from the same set. Note there were two additional mini-blocks of 1D trials only, at the end of first- and at the start of the second phase (during anatomical scan, see below). The replication sample completed the same procedure with the same break length, but without MRI scanning. That is, both phases were completed in a behavioral testing room. Details of each task and the stimuli are described below. Behavioral data was recorded during all experiment phases. MRI data was recorded during phase 2. We additionally collected eye-tracking data (EyeLink 1000; SR Research Ltd.; Ottawa, Canada) both during the staircasing and the main decision making task to ensure continued fixation (data not presented). The overall experiment lasted XXX minutes on average.

Room, Luminance and Apparatus Behavioral sessions were conducted in a dimly lit room without natural light sources, such that light fluctuations could not influence the perception of the features. A small lamp was stationed in the corner of the room, positioned so it would not cast shadows on the screen. The lamp had a light bulb with 100% color rendering index, i.e. avoiding any influence on color perception. Participants sat on a height adjustable chair at a distance of 60 cm from a 52 cm horizontally wide, Dell monitor (resolution: 1920 x 1200, refresh rate 1/60 frames per second). Distance from the monitor was fixed using a chin-rest with a head-bar. Stimuli were presented using psychtoolbox version 3.0.11 [49–51] in MATLAB R2017b [52]In the MRI-scanner room lights were switched off and light sources in the operating room were covered in order to prevent interference with color perception or shadows cast on the screen. Participants lay inside the scanner at distance of 91 cm from a 27 cm horizontally wide screen on which the task was presented a D-ILA JVC projector (D-ILa Projektor SXGA, resolution: 1024x768 , refresh rate: 1/60 frames per second). Stimuli were presented using psychtoolbox version 3.0.11 [49–51] in MATLAB R2012b [53] on a Dell precision T3500 computer running windows XP version 2002.

Stimuli Each cloud of dots was presented on the screen in a circular array with 7° visual angle in diameter. In all trials involving two clouds, the clouds appeared with 4° visual angle distance between them, including a fixation circle (2° diameter) in the middle, resulting in a total of 18° field of view [following total apparatus size from 38]. Each cloud consisted of 48 square dots of 3x3 pixels. We used four specific motion and four specific color features.

To prevent any bias resulting from the correspondence between response side and dot motion, each of the four motion features was constructed of two angular directions rotated by 180°, such that motion features reflected an axis of motion, rather than a direction. Specifically, we used the four combinations: 0°-180° (left-right), 45°-225° (bottom right to upper left), 90°-270° (up-down) and 135°-315° (bottom left - upper right). We used a Brownian motion algorithm [e.g. 38], meaning in each frame a different set of given amount of coherent dots was chosen to move coherently in the designated directions in a fixed speed, while the remaining dots moved in a random direction (Fig. S1). Dots speed was set to 5° per second [i.e. 2/3 of the aperture diameter per second, following 38]. Dots lifetime was not limited. When a dot reached the end of the aperture space, it was sent 'back to start', i.e. back to the other end of the aperture. Crucially, the number of coherent dots (henceforth: motion-coherence) was adjusted for each participant throughout the staircasing procedure, starting at 0.7 to ensure high accuracy [see 38]. An additional type of motion-direction was 'random-motion' and was used in 1D color clouds. In these clouds, dots were split to 4 groups of 12, each assigned with one of the four motion features and their adjusted-coherence level, resulting in a balanced subject-specific representation of random motion.

In order to keep the luminance fixed, all colors presented in the experiment were taken from the YCbCr color space with a fixed luminance of Y = 0.5. YCbCr is believed to represent human perception in a relatively accurate manner [cf. 54]. In order to generate an adjustable parameter for the purpose of staircasing, we simulated a squared slice of the space for Y = 0.5 (Fig. S1) in which the representation of the dots color moved using a Brownian motion algorithm as well. Specifically, all dots started close to the (gray) middle of the color space, in each frame a different set of 30% of dots was chosen to move coherently towards the target color in a certain speed whereas all the rest were assigned with a random direction. Perceptually, this resulted in all the dots being gray at the start of the trial and slowly taking on the designated color. Starting point for each color was chosen based on pilot studies and was set to a distance of 0.03-0.05 units in color space from the middle. Initial speed in color space (henceforth: color-speed) was set so the dots arrive to their target (23.75% the distance to the corner from the center) by the end of the stimulus presentation (1.6s). i.e. distance to target divided by the number of frames per trial duration. Color-speed was adjusted throughout the staircasing procedure. An additional type of color was 'no color' for motion 1D trials for which we used the gray middle of the color space.

Staircasing task In order to ensure RTs mainly depended on associated values and not on other stimulus properties (e.g. salience), we created a staircasing procedure that was conducted *prior to value learning*. In this procedure, motion-coherence and color-speed were adjusted for each participant in order to minimize between-feature detection time differences. As can be seen in Fig. S1, in this perceptual detection task participants were cued (0.5s) with either a small arrow (length 2°) or a small colored circle (0.5° diameter) to indicate which motion-direction or color they should choose in the upcoming decision. After a short gray (middle of YCbCr) fixation circle (1.5s, diameter 0.5°), participants made a decision between the two clouds (1.6s). Clouds in this part could be either both single-feature or both dual-features. In dual feature trials, each stimulus had one color and one motion feature, but the cue indicated either a specific motion or a specific color. After a choice, participants received feedback (0.4s) whether they were (a) correct and faster than 1 second, (b) correct and slower or (c) wrong. After a short fixation (0.4s), another trial started. All timings were fixed in this part. Participants were instructed to always look at the fixation circle in the middle of the screen throughout this and all subsequent tasks. To motivate participants and continued perceptual improvements during the later (reward related) task-stages, participants were told that if they were correct and faster than 1 second in at least 80% of the trials, they will receive an additional monetary bonus of 2 Euros.

The staircasing started after a short training (choosing correct in 8 out of 12 consecutive trials mixed of both contexts) and consisted of two parts: two *adjustment blocks* an two *measurement blocks*. All adjustments of color-speed and motion-coherence followed this formula:

$$\theta_i^{t+1} = \theta_i^t + \alpha\theta_i^t \frac{\overline{RT_i^t} - RT^0}{RT^0} \tag{1}$$

where θ_i^{t+1} represents the new coherence/speed for motion or color feature i during the upcoming time interval/block $t+1$, θ_i^t is the level at the time of adjustment, $\overline{RT_i^t}$ is the mean RT for the specific feature i during time interval t, RT_0 is the "anchor" RT towards which the adjustment is made and α represents a step size of the adjustment, which changed over time as described below.

The basic building block of *adjustment blocks* consisted of 24 cued-feature choices for each context ($4 \times 3 \times 2 = 24$, i.e. 4 colors, each discriminated against 3 other colors, on 2 sides of screen). The same feature was not cued more than twice in a row. Due to time constrains, we could not include all possible feature-pairing combinations between the cued and uncued features. We therefore pseudo-randomly choose from all possible background combinations for each feature choice (unlike later stages, this procedure was validated on and therefore included also trials with identical background features). In the first adjustment block, participants completed 72 trials, i.e. 36 color-cued and 36 motion-cued, interleaved in chunks of 4-6 trials in a non-predictive manner. This included, for each context, a mixture of one building block of 2D trials and half a block of 1D trials, balanced to include 3 trials for each cued-feature. 1D or 2D trials did not repeat more than 3 times in a row. At the end of the first adjustment block, the mean RT of the last 48 (accurate) trials was taken as the anchor (RT^0) and each individual feature was adjusted using the above formula with $\alpha = 1$. The second adjustment block started with 24 motion-cued only trials which were used to compute a new anchor. Then, throughout a series of 144 trials (72 motion-cued followed by 72 color-cued trials, all 2D), every three correct answers for the same feature resulted in an adjustment step for that specific feature (Eq. 1) using the average RT of these trials ($\overline{RT_i^t}$) and the motion anchor RT^0 for both contexts. This resulted in a maximum of six adjustment steps per feature, where alpha decreased from 0.6 to 0.1 in steps of 0.1 to prevent over-adjustment.

Next, participants completed two *measurement blocks* identical in structure to the main task (see below) with two exceptions: First, although this was prior to learning the values, they were perceptually cued to chose the feature that later would be assigned with the highest value. Second, to keep the relevance of the feature that later would take the lowest value (i.e. would rarely be chosen), we added 36 additional trials cued to choose that feature (18 motion and 18 color trials per block).

Outcome learning task After the staircasing and prior to the main task, participants learned to associate each feature with a deterministic outcome. Outcomes associated with the four features on each contexts were 10, 30, 50 and 70 credit-points. The value mapping to perceptual features was assigned randomly between participants, such that all possible color- and all possible motion-combinations were used at least once ($4! = 24$ combinations per context). We excluded motion value-mapping that correspond to clockwise or counter-clockwise ordering. The outcome learning task consisted only of single-feature clouds, i.e. clouds without coherent motion or dots 'without' color (gray). Therefore each cloud in this part only represented a single feature. To encourage mapping of the values for each context on similar scales, the two clouds could be either of the same context (e.g. color and color) or from different contexts (e.g. color and motion). Such context-mixed trials did not repeat in other parts of the experiment.

The first block of the outcome learning task had 80 *forced choice* trials (5 repetitions of 16 trials: 4 values $\times$ 2 Context $\times$ 2 sides of screen), in which only one cloud was presented, but participants still had to choose it to observe its associated reward. These were followed by mixed blocks of 72 trials which included 16 *forced choice* interleaved with 48 *free choice* trials between two 1D clouds (6 value-choices: 10 vs 30/50/70, 30 vs 50/70, 50 vs 70 $\times$ 4 context combinations $\times$ 2 sides of screen for highest value). To balance the frequencies with which feature-outcome pairs would be chosen, we added 8 forced choice trials in which choosing the lowest value was required. Trials were pseudo-randomized so no value would repeat more than 3 times on the same side and same side would not be chosen more the three consecutive times. Mixed blocks repeated until participants reached at least 85% accuracy of choosing the higher valued cloud in a block, with a minimum of two and a maximum of four blocks. Since all clouds were 1D and choice could be between contexts, these trials started without a cue, directly with the presentation of two 1D clouds (1.6s). Participants then made a choice, and after short fixation (0.2s) were presented with the value of both chosen and unchosen clouds (0.4s, with value of choice marked with a square around it, see Fig. S1). After another short fixation (0.4s) the next trial started. Participants did not collect reward points in this stage, but were told that better learning of the associations will result in more points, and therefore more money later. Specifically, in the MRI experiment participants were instructed that credit points during the main task will be converted into a monetary bonus such that every 600 points they will receive 1 Euro at the end. The behavioral replication cohort received 1 Euro for every 850 points.

Main task preparation In preparation of the main task, participants performed one block of 1D trials at the end of phase 1 and then at the start of the MRI session during the anatomical scan. These blocks were included to validate that changing presentation mediums between phases (computer screen versus projector) did not introduce a perceptual bias to any features and as a final correction for post value-learning RT differences between contexts. Each block consisted of 30 color and 30 motion 1D trials interleaved in chunks of 4-7 trials in a non-predictive manner. The value difference between the clouds was fixed to 20 points (10 repetitions of 3 value comparisons $\times$ 2 contexts). Trials were pseudo-randomized so no target value was repeated more than once within context (i.e. not more than twice all in all) and was not presented on the same side of screen more than 3 consecutive trials within context and 4 in total. In each trial, they were first presented with a contextual cue (0.6s) for the trial, followed by short fixation (0.5s) and the presentation of two single-feature clouds of the cued context (1.6s) and had to choose the highest valued cloud. After a short fixation (0.4s), participants were presented with the chosen cloud's outcome (0.4s). The timing of the trials was fixed and shorter than in the remaining main task because no functional MRI data was acquired during these blocks. Participants were instructed that from the first preparation block they started to collect the rewards. Data from these 1D block were used to inspect and adjust for potential differences between the MRI and the behavior setup. First, participants reacted generally slower in the scanner ($t(239) = -9.415$, $p < .001$, paired t-test per subject per feature). Importantly, however, we confirmed that this slowing was uniform across features, i.e. no evidence was found for a specific feature having more RT increase than the rest (ANOVA test on the difference between the phases, $F(7, 232) = 1.007$, $p = .427$). Second, because pilot data indicated increased RT differences between contexts after the outcome learning task we took the mean RT difference between color and motion trials in the second mini-block in units of frames (RT difference divided by the refresh rate), and moved the starting point of each color relative to their target color, the number of frames $\times$ its speed. Crucially, the direction of the move (closer/further to target) was the same for all colors, thus ensuring not to induce within-context RT differences.

Main task Finally, participants began with the main experiment inside the scanner. Participants were asked to choose the higher-valued of two simultaneously presented random dot kinematograms, based on the previously learned feature-outcome associations. As described in the main text, each trial started with a cue that indicated the current task context (color or motion). In addition, both clouds could either have two features (each a color and a motion, *2D trials*) or one feature only from the cued context (e.g., colored, but randomly moving dots).

The main task consisted of four blocks in which 1D and 2D trial were intermixed. Each block contained 36 1D trials (3 EV $\times$ 2 Contexts $\times$ 6 repetitions) and 72 2D trials (3 EV $\times$ 2 Contexts $\times$ 12 feature-combinations, see fig1c). Since this task took part in the MRI, the duration of the fixation circles were drawn from an truncated exponential distribution with a mean of μ=0.6s (range 0.5s-2.5s) for the interval between cue and stimulus, a mean of μ=3.4s (1.5s-9s) for the interval between stimulus and outcome and a mean of μ=1.25s (0.7s-6s) for the interval between outcome and the cue of the next trial. The cue, stimulus and outcome were presented for 0.6s, 1.6sand 0.8s, respectively. Timing was optimized using VIF-calculations of trial-wise regression models (see Classification procedure section below).

The order of trials within blocks was controlled as follows: the cued context stayed the same for 4-7 trials (in a non-predictive manner), to prevent context confusion caused by frequent switching. No more than 3 repetitions of 1D or 2D trials within each context could occur, and no more than 5 repetition overall. The target did not appear on the same side of the screen on more than 4 consecutive trials. Congruent or incongruent trials did not repeat more than 3 times in a row. In order to avoid repetition suppression, i.e. a decrease in the fMRI signal due to a repetition of information [e.g. 55, 56], no target feature was repeated two trials in a row, meaning the EV could repeat maximum once (i.e. one color and one motion). As an additional control over repetition, we generated 1000 designs according the above-mentioned rules and choose the designs in which the target value was repeated in no more than 10% of trials across trial types, as well as when considering congruent, incongruent or 1D trials separately.

Behavioral analysis

RT data was analyzed in R (R version 3.6.3 [57], RStudio version 1.3.959 [58]) using linear mixed effect models (lmer in lme4 1.1-21: [59]). When describing main effects of models, the χ^2 represents Type II Wald χ^2 tests, whereas when describing model comparison, the χ^2 represents the log-likelihood ratio test. Model comparison throughout the paper was done using the 'anova' function. Regressors were scaled prior to fitting the models for all analyses. The behavioral model that we found to fit the behavioral RT data best was:

$$\begin{aligned}\log RT_k^t = \beta_0 + \gamma_{0k} + \beta_1 EV + \beta_2 Congruency_t + \beta_3 Congruency_t \times EV_{back_t} + \beta_4 Congruency_t \times EV_t \\ + \nu_1 t + \nu_2 side_t + \nu_3 switch_t + \nu_4 context_t\end{aligned} \tag{2}$$

where $\log RT_k^t$ is the log reaction time of subject k in trial t, β_0 and γ_{0k} represent global and subject-specific intercepts, ν-coefficients reflect nuisance regressors ($side$ of target object, trials since last context $switch$ and the current $context$), β_1 to β_4 captured the fixed effect of EV, Congruency, Congruency $\times$ EV$_{back}$ and Congruency $\times$ EV, respectively. The additional models reported in the SI included intercept terms specific for each factor level, nested within subject (for EV, Block and Context, see Fig. S2). Investigations of alternative parametrizations of the values can be found in Fig. S3.

Accuracy data was analyzed in R (R version 3.6.3 [57], RStudio version 1.3.959 [58]) using generalized linear mixed effect models (glmer in lme4 1.1-21: [59]) employing a binomial distribution family with a 'logit' link function. Regressors were scaled prior to fitting the models for all analyses. No-answer trials of were excluded from this analysis. The model found to fit the behavioral accuracy data best was almost equivalent to the RT model, except for the fourth term involving Congruency $\times$ switch:

$$\begin{aligned}ACC_k^t = \beta_0 + \gamma_{0k} + \beta_1 EV + \beta_2 Congruency_t + \beta_3 Congruency_t \times EV_{back_t} + \beta_4 Congruency_t \times switch_t \\ + \nu_1 t + \nu_2 side_t + \nu_3 switch_t + \nu_4 context_t\end{aligned} \tag{3}$$

where ACC_k^t is the accuracy (1 for correct and 0 for incorrect) of subject k in trial t and all the rest of the regressors are equivalent to Eq. 2. We note that the interaction Congruency $\times$ switch indicates that participants were more accurate the further they were from a context switch point.

fMRI data

fMRI data acquisition MRI data was acquired using a 32-channel head coil on a research-dedicated 3-Tesla Siemens Magnetom TrioTim MRI scanner (Siemens, Erlangen, Germany) located at the Max Planck Institute for Human Development in Berlin, Germany. High-resolution T1-weighted (T1w) anatomical Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequences were obtained from each participant to allow registration and brain surface reconstruction (sequence specification: 256 slices; TR = 1900 ms; TE = 2.52 ms; FA = 9 degrees; inversion time (TI) = 900 ms; matrix size = 192 x 256; FOV = 192 x 256 mm; voxel size = 1 x 1 x 1 mm). This was followed with two short acquisitions with six volumes each that were collected using the same sequence parameters as for the functional scans but with varying phase encoding polarities, resulting in pairs of images with distortions going in opposite directions between the two acquisitions (also known as the blip-up / blip-down technique). From these pairs the displacements were estimated and used to correct for geometric distortions due to susceptibility-induced field inhomogeneities as implemented in the the fMRIPrep preprocessing pipeline. In addition, a whole-brain spoiled gradient recalled (GR) field map with dual echo-time images (sequence specification: 36 slices; A-P phase encoding direction; TR = 400 ms; TE1 = 4.92 ms; TE2 = 7.38 ms; FA = 60 degrees; matrix size = 64 x 64; 619 FOV = 192 x 192 mm; voxel size = 3 x 3 x 3.75 mm) was obtained as a potential alternative to the method described above. However, this GR frield map was not used in the preprocessing pipeline. Lastly, four functional runs using a multi-band sequence (sequence specification: 64 slices in interleaved ascending order; anterior-to-posterior (A-P) phase encoding direction; TR = 1250 ms; echo time (TE) = 26 ms; voxel size = 2 x 2 x 2 mm; matrix = 96 x 96; field of view (FOV) = 192 x 192 mm; flip angle (FA) = 71 degrees; distance factor = 0, MB acceleration factor = 4). A tilt angle of 30 degrees from AC-PC was used in order to maximize signal from the orbitofrontal cortex (OFC, see [60]). For each functional run, the task began after the acquisition of the first four volumes (i.e., after 5.00 s) to avoid partial saturation effects and allow for scanner equilibrium. Each run was about 15 minutes in length, including a 20 seconds break in the middle of the block (while the scanner is running) to allow participants a short break. We measured respiration and pulse during each scanning session using pulse oximetry and a pneumatic respiration belt part of the Siemens Physiological Measurement Unit.

BIDS conversion and defacing Data was arranged according to the brain imaging data structure (BIDS) specification [61] using the HeuDiConv tool (version 0.6.0.dev1; freely available from https://github.com/nipy/heudiconv). Dicoms were converted to the NIfTI-1 format using dcm2niix [version 1.0.20190410 GCC6.3.0; [62]]. In order to make identification of study participants highly unlikely, we eliminated facial features from all high-resolution structural images using pydeface (version 2.0; available from https://github.com/poldracklab/pydeface). The data quality of all functional and structural acquisitions were evaluated using the automated quality assessment tool MRIQC [for details, [see 63], and the MRIQC documentation]. The visual group-level reports confirmed that the overall MRI signal quality was consistent across participants and runs.

fMRI preprocessing Data was preprocessed using *fMRIPrep* 1.2.6 (Esteban et al. [64]; Esteban et al. [65]; RRID:SCR_016216), which is based on *Nipype* 1.1.7 (Gorgolewski et al. [66]; Gorgolewski et al. [67]; RRID:SCR_002502). Many internal operations of *fMRIPrep* use *Nilearn* 0.5.0 [68, RRID:SCR_001362], mostly within the functional processing workflow.

Specifically, the T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) using `N4BiasFieldCorrection` [69, ANTs 2.2.0], and used as a T1w-reference throughout the workflow. The anatomical image was skull-stripped using `antsBrainExtraction.sh` (ANTs 2.2.0), using OASIS as the target template. Brain surfaces were reconstructed using `recon-all` [FreeSurfer 6.0.1, RRID:SCR_001847, 48], and the brain masks were estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle [RRID:SCR_002438, 47]. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c [70, RRID:SCR_008796] was performed through nonlinear registration with `antsRegistration` [ANTs 2.2.0, RRID:SCR_004757, 71], using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using `fast` [FSL 5.0.9, RRID:SCR_002823, 72].

To preprocess the functional data, a reference volume for each run and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on two echo-planar imaging (EPI) references with opposing phase-encoding directions, using `3dQwarp` [73] (AFNI 20160207). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using `bbregister` (FreeSurfer), which implements boundary-based registration [74]. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` [FSL 5.0.9, 75]. BOLD runs were slice-time corrected using `3dTshift` from AFNI 20160207 [73, RRID:SCR_005927] and aligned to the middle of each TR. The BOLD time-series (including slice-timing correction) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*.

Several confound regressors were calculated were calculated during preprocessing: Six head-motion estimates (see above), Framewise displacement, six anatomical component-based noise correction components (aCompCorr) and 18 physiological parameters (8 respiratory, 6 heart rate and 4 of their interaction). The head-motion estimates were calculated during motion correction (see above). Framewise displacement was calculated for each functional run, using the implementations in *Nipype* [following the definitions by 76]. A set of physiological regressors were extracted to allow for component-based noise correction [*CompCor*, 77]. Principal components are estimated after high-pass filtering the BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor, unused) and anatomical (aCompCor). For aCompCor, six components are calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction, and co-registrations to anatomical and template spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels [78]. Lastly, for the 18 physiological parameters, correction for physiological noise was performed via RETROICOR [79, 80] using Fourier expansions of different order for the estimated phases of cardiac pulsation (3rd order), respiration (4th order) and cardio-respiratory interactions (1st order) [81]: The corresponding confound regressors were created using the Matlab PhysIO Toolbox ([82], open source code available as part of the TAPAS software collection: `https://www.translationalneuromodeling.org/tapas`. For more details of the pipeline, and details on other confounds generated but not used in our analyses, see the section corresponding to workflows in *fMRIPrep*'s documentation.

For univariate analyses, BOLD time-series were re-sampled to MNI152NLin2009cAsym standard space in the *fMRIPrep* pipeline and then smoothed using SPM [83, SPM12 (7771)] with 8mm FWHM, except for ROI generation, where a 4mm FWHM kernel was used. Multivariate analyses were conducted in native space, and data was smoothed with 4mm FWHM using SPM [83, SPM12 (7771)]. Classification analyses further involved three preprocessing steps of voxel time-series: First, extreme-values more than 8 standard deviations from a voxels mean were corrected by moving them by 50% their distance from the mean towards the mean (this was done to not bias the last z scoring step). Second, the time-series of each voxel was detrended, a high-pass filter at 128 Hz was applied and confounds were regressed out in one action using *Nilearn* 0.6.2 [68]. Lastly, the time-series of each voxel for each block was z scored.

Univariate fMRI analysis

All GLMs were conducted using SPM12 [83, SPM12 (7771)] in MATLAB [52]. All GLMs consisted of two regressors of interest corresponding to the onsets of the two trial-types (1D/2D, except for one GLM where 2D onsets were split by Congruency) and included one parametric modulator of EV assigned to 1D onset and different combinations of parametric modulators of EV, Congruency, EV_{back} and their interactions (see Fig. S9 for GLM visualization). All parametric modulators were demeaned before entering the GLM, but not orthogonalized. Regressors of no interest reflected cue onsets in Motion and Color trials, stimulus onsets in wrong and no-answer trials, outcome onsets and 31 nuisance regressors (e.g. motion and physiological parameters, see fMRI-preprocessing). The duration of stimulus regressors corresponded to the time the stimuli were on screen. The durations for the rest of the onset regressors were set to 0. Microtime resultion was set to 16 (64 slices / 4 MB factor) and microtime onset was set to the 8 (since slice time correction aligned to middle slice, see fMRI-preprocessing). Data for all univariate analyses were masked with a whole brain mask computed as intercept of each functional run mask generated from fMRIprep [47, 48]. MNI coordinates were translated to their corresponding brain regions using the automated anatomical parcellation toolbox [84–86, AAL3v1] for SPM. We verified the estimability of the design matrices by assessing the Variance Inflation Factor (VIF) for each onset regressor in the HRF-convolved design matrix. Specifically, for each subject, we computed the VIF (assisted by scripts from `https://github.com/sjgershm/ccnl-fmri`) for each regressor in the HRF-convolved design matrix and averaged the VIFs of corresponding onsets across the blocks. None of the VIFs surpassed a value of 3.5 (a value of 5 is considered a conservative indicator for overly colinear regressors, e.g. [87], see Fig.S9 for details). Detailed descriptions of all GLMs are reported in the main text. Additional GLMs verifying the lack of Congruency in any frontal region can be found in Fig.S9.

vmPFC functional ROI In order to generate a functional ROI corresponding to the vmPFC in a reasonable size, we re-ran the GLM with only EV modulators (i.e. this GLM had no information regarding the contextually irrelevant context) on data that was smoothed at 4mm. We then threshold the EV contrasts for 1D and 2D trials (EV_{1D} + EV_{2D} >0) at $p < .0005$. The group ROI was generated in MNI space and included 998 voxels. Multivariate analyses were conducted in native space and the ROI was transformed to native space using ANTs and nearest neighbor interpolation [ANTs 2.2.0 71] while keeping only voxels within the union of subject- and run-specific brain masks produced by the fMRIprep pipeline [47, 48]. The resulting subject-specific ROIs therefore had varying number of voxels (μ = 768.14, σ = 65.62, min = 667, max = 954).

Multivariate analysis

Classification procedure The training set for all analyses consisted of fMRI data from behaviorally accurate 1D trials. For each trial, we took the TR corresponding to approx. 5 seconds after stimulus onset ($round(onset + 5)$) to match the peak of the Haemodynamic Response Function (HRF) estimated by SPM [83]. Classification training was done using a leave-one-run-out scheme across the four runs with 1D trials. To avoid bias in the training set after sub-setting only to behaviorally accurate trials (i.e. over-representation of some information) we up-sampled each training set to ensure equal number of examples in the training set for each combination of EV (3), Context (2) and Chosen-Side (2). Specifically, if one particular category was less frequent than another (e.g., more value-30, left, color trials than value-50, left-color trials) we up-sampled that example category by randomly selecting a trial from the same category to duplicate in the training set, whilst prioritising block-wise balance (i.e., if one block had 2 trials in the chunk and another block had only 1, we first duplicated the trial from under-represented block etc.). We did not up-sample the testing set. Decoding was conducted using multinomial logistic regression as implemented in *scikit-learn* 0.22.2 [88] set to multinomial (in opposed to one-vs-all) with C-parameter 1.0, lbgfs solver with a 'l2' penalty for regularization. The classifier provided for each trial in the testing block one probability (or: predicted probability) per class that was given to it. To avoid bias in the modeling of the classifier's predictions (i.e. one probability for each class) we performed outlier-correction, i.e. rounded up values smaller than 0.00001 and down values bigger than 0.99999. Due to technical reasons, we averaged the classifier probabilities across the nuisance effects, i.e. obtaining one average probability for each combination of relevant and irrelevant values. This resulted in 36 probabilities per participant, one for each combination of EV level (three levels), irrelevant value of the chosen side and irrelevant value of the non-chosen side (12 combinations, see Fig. 1). Note that the relevant value of the unchosen cloud was always EV - 20 and therefore we did not include this as a parameter of interest. After averaging, we computed for each combination of values the EV_{back}, Congruency and alternative parameters (see Fig. S8). The main model comparison, as well as the lack of effects of any nuisance regressor, was confirmed on a dataset with raw, i.e. non-averaged, probabilities (see Fig S6 and S8). Throughout all the analyses, each regressor was scaled prior to fitting the models. Lastly, for the analysis of $\text{P}_{\text{EV}_{\text{back}}}$ (Fig. 5d.) and for Fig. 7 we also included behaviorally wrong trials.

Verifying design trial-wise estimability To verify that the individual trials are estimatable and as a control over multi-colinearity [87], we convolved a design matrix with the HRF for each subject with one regressor per stimuli (432 regressors with duration equal to the stimulus duration), two regressor across all cues (split by context) and three regressor for all outcomes (one for each EV). We then computed the VIF for each stimulus regressor (i.e. how predictive is each regressor by the other ones). None of the VIFs surpassed 1.57 across all trials and subjects ($\mu_{VIF} = 1.42, \sigma_{VIF} = .033$, min = 1.34). When repeating this analysis with a GLM in which also outcomes were split into trialwise regressors, we found no stimuli VIF larger than 3.09 ($\mu_{VIF} = 2.64, \sigma_{VIF} = .132$, min = 1.9). Note that 1 is the minimum (best) value and 5 is a relatively conservative threshold for colinearity issues ([e.g. 87]). This means that the BOLD responses of individual trials can be modeled separately and should not have colinearity issues with other stimuli nor with the outcome presentation of each trial.

Modelling class probabilities The classifier provided one probability to each class, given the data (all probabilities for each trial sum to 1). Probabilities were analyzed in R (R version 3.6.3 [57], RStudio version 1.3.959 [58]) with Generalized Linear Mixed Models using Template Model Builder (glmmTMB, [89]) models, employing a beta distribution family with a 'logit' link function. When describing main effects of models, the χ^2 represents Type II Wald χ^2 tests, whereas when describing model comparison, the χ^2 represents the log-likelihood ratio test. Model comparison throughout the paper was done using the 'anova' function.

The value similarity analyses asked whether the predicted probabilities reflected the difference from the objective probability class. The model we found to best explain the data was:

$$P^k_{t,c} = \beta_0 + \gamma_{0k} + \beta_1|EV_t - Class_{c,t}| + \beta_2|EV_t - Class_{c,t}|EV_{back_t} \tag{4}$$

where $\text{P}^k_{t,c}$ is the probability assigned to class c in trial t for subject k, β_0 and γ_{0k} represent global and subject-specific intercepts, $|EV_t - Class_{c,t}|$ is the absolute difference between the EV of the trial and the class the probability is assigned to and $|EV_t - Class_{c,t}|EV_{back_t}$ is the interaction of this absolute difference with EV_{back}. For models nested in the levels of EV, we included $\zeta_{0_{k_v}}$, which is the EV-specific intercept nested within each within each subject level.

For the feature similarity model we substituted $|EV_t - Class_{c,t}|$ with a "similarity" parameter that encoded the perceptual similarity between each trial in the test set and the perceptual features that constituted the training examples of each class of the classifier. For 1D trials, this perceptual parameter was identical to the value similarity parameter ($|EV_t - Class_{c,t}|$). This was because from the shown pairs of colors, both colors overlapped between training and test if the values were identical; one color overlapped if the values were different by one reward level (e.g. a 30 vs 50 comparison corresponded to two trials that involved pink vs green and green vs orange, i.e. sharing the color green); and no colors overlapped if the values were different by two levels (30 vs 70). On 2D trials however, due to changing background features and their value-difference variation, perceptual similarity of training and test was not identical to value similarity. Even though both the value similarity and the perceptual similarity parameter correlated ($\rho = .789$, $\sigma = .005$), we found that the value similarity model provided a better AIC score (value similarity AIC: -3898, Feature similarity AIC: -3893, Fig. 4). Detailed description with examples can be found in Fig. S6. Crucially, even when keeping the value difference of the irrelevant features at 20, thus limiting the testing set only to trials with feature-pairs that were included in the training, our value similarity model provided a better AIC (-1959) than the feature similarity model (-1956). To test for a perceptual alternative of $\mathrm{EV}_{\mathrm{back}}$ we substituted the corresponding parameter from the model with $\mathrm{Similarity}_{\mathrm{back}}$. This perceptual parameter takes on 1 if the perceptual feature corresponding to the $\mathrm{EV}_{\mathrm{back}}$ appeared in the 1D training class (as highest or lowest value) and 0 otherwise. As described in the main text, none of the perceptual-similarity encoding alternatives provided a better fit than our models that focused on the expected values the features represented.

When modelling the probability of the objective EV, the model we found to explained the data best was:

$$P^k_{t,EV} = \beta_0 + \gamma_{0k} + \beta_1 EV_{back_t} \tag{5}$$

where $\mathrm{P}^k_{t,EV}$ is the probability assigned to the objective class (corresponding to EV of the trial t) for subject k, β_0 and γ_{0k} represent global and subject-specific intercepts and EV_{back} is the maximum of the two ignored values (or the EV of the contextually irrelevant context). For models nested in the levels of EV, we included $\zeta_{0_{kv}}$ which is EV specific intercept nested within each within each subject level (see Fig. S8). Investigations of alternative parametrizations of the values can be found in Fig. S8.

When modelling the probability of $\mathrm{EV}_{\mathrm{back}}$, we did not average across nuisance regressors. Our baseline model was: $P^k_{t,EV_{back}} = \beta_0 + \gamma_{0k} + \nu_1 side(t) + \nu_2 switch(t) + \nu_3 context(t)$. Neither including a main effect nor interactions between EV, $\mathrm{EV}_{\mathrm{back}}$ and Congruency improved model fit. When including behaviorally wrong trials in the model, we used `drop1` in combination with χ^2-tests from `lmer4` package [59] to test which of the main effects or interactions improves the fit. This resulted in the following model as best explaining the data:

$$\begin{aligned} P^k_{t,EV_{back}} = \beta_0 + \gamma_{0k} + \beta_1 EV_t \times EV_{back_t} + \beta_2 Congruency_t \times Accuracy_t \\ + \nu_1 t + \nu_2 side_t + \nu_3 switch_t + \nu_4 context_t \end{aligned} \tag{6}$$

where $\mathrm{P}^k_{t,EV_{back}}$ is the probability assigned to the EV_{back} class (corresponding to EV_{back} of trial t) for subject k, β_0 and γ_{0k} represent global and subject-specific intercepts, EV is the maximum of the two relevant and EV_{back} is the maximum of the two ignored values. Congruency reflects whether the actions chosen in the relevant vs. irrelevant context would be the same, and the Accuracy regressor has 1 if participants chose the highest relevant value and 0 otherwise. We note that the interaction EV $\times$ EV_{back} ($\chi^2_{(1)} = 4.18$, $p = .041$) indicates higher in trials in which EV and EV_{back} were more similar, the probability assigned to EV_{back} was higher. However, we find this effect hard to interpret since this corresponds to the value similarity effect we previously reported.

Parallel representation of outcomes in vmPFC. To compute the correlations between each pair of classes we transformed the probabilities for each class using a multinomial logit transform. For example, for class 30 we performed probabilities were transformed with $mlogit(P_{t,30}) = 0.5(\log \frac{P_{t,30}}{P_{t,50}} + \log \frac{P_{t,30}}{P_{t,70}})$. To examine the relationship between EV and $\mathrm{EV}_{\mathrm{back}}$, we only included 2D trials in which EV $\neq$ $\mathrm{EV}_{\mathrm{back}}$. This allowed us to categorize all three probabilities as either EV, $\mathrm{EV}_{\mathrm{back}}$ or Other, whereby Other reflected the value that was neither the EV, nor the $\mathrm{EV}_{\mathrm{back}}$. To prevent bias we included only trials in which Other was presented on screen (as relevant or irrelevant value). We then averaged across nuisance regressors (see Classification procedure) and computed the correlation across all trials. Lastly, we Fisher z-transformed the correlations ($0.5 \log \frac{1+\rho}{1-\rho}$) to approximate normality for the t test. To validate these results, we performed an additional model comparison in which we added a term of the logit transformed $P_{EV_{back}}$ or of P_{other} to Eq. 5 ($\beta_2 mlogit(P_{t,EV_{back}})$ or $\beta_2 mlogit(P_{t,Other})$,respectively). As reported in the main text, adding a term reflecting $P_{EV_{back}}$ resulted in a smaller (better) AIC score than when we added a term for P_{other} (-567,-475, respectively). This was also preserved when running the analysis including nuisance regressors (see νs in Eq. 2) on the non-averaged data (AICs: -5913.3,-5813.3). We note that subsetting the data the way we did resulted in a strong negative correlation in the design matrix between EV and EV_{back} ($\rho = -0.798$, averaged across subjects). Although this should not directly influence our interpretation, we validated the results by using alternative models with effects hierarchically nested within the levels of EV *and* EV_{back} (Averaged data AICs: -560, -463, Raw data AICs: -5906.8,-5804.3)

Linking MRI effects to behavior We showed that subjects who had a stronger effect of Congruency on their RT also had a stronger effect of EV_{back} on P_{EV}, as well as a stronger correlation between P_{EV} and $P_{\text{EV}_{\text{back}}}$.

The model used to obtain subject-specific Congruency and Congruency x EV_{back} slopes was:

$$\begin{aligned} \log RT_t^k = \beta_0 + \gamma_{0k} + \beta_1 EV + \beta_2 Congruency_t + \beta_3 Congruency_t EV_{back_t} \\ + \gamma_{1k} Congruency + \gamma_{2k} Congruency + \gamma_{3k} EV_{back_t} \\ + \nu_1 t + \nu_2 side_t + \nu_3 switch_t + \nu_4 context_t \end{aligned} \tag{7}$$

where all the notations are the same as in Eq. 2. γ_{1k} represents the subject-specific slope for Congruency for subject k and γ_{2k} for the interaction of Congruency and EV_{back}.

To extract subject-specific slopes for the effect of EV_{back} on P_{EV} we included a term for this effect ($\gamma_{1k} EV_{back_t}$) in Eq. 5. Due to model convergence issues, the we had to drop the subject-specific intercept (γ_{0k}) in that model.

For the correlation of P_{EV} and $P_{EV_{back}}$ we only used trials in which EV $\neq$ EV_{back}. Probabilities were first multinomial logit and then Fisher z-transformed (see above) and averaged across trials to achieve one correlation value per subject. In the main text and in Fig 5 we did not average the data to achieve maximum sensitivity to trial-wise variations. The results reported in the main text replicate when running the same procedure while averaging the data across nuisance regressors following the multinomial logit transformation ($R = .38$, $p = .023$).

Context decoding Classification of task context followed the same procedures as when decoding of EV (see 'Classification procedure'), albeit the classes given to the classifier for each 1D train example were the context, i.e. 'Color' or 'Motion'. Up-sampling was done in the same manner, resulting in 4 training sets that are each balanced across EV, Context and Side of target object, and balanced block-wise as much as possible.

To perform the analysis shown in Fig. 6d, we included a main effect of $P_{context}$ in Eq. 5 that was logit-transformed ($logit(P) = \log \frac{P}{1-P}$) and scaled for each subject, thus adding the term $\beta_2 logit(P_{Context})$. Note that since there are only 2 classes, there is no need for multinomial logit transformation.

Neural representations of EV, EV_{back} and Context as predictors of behavioral accuracy We used hierarchical model comparison to directly test the influence of neural representation of EV, EV_{back} and Context on behavioral accuracy separately for congruent and incongruent trials. First, we tested if adding $logit(P_{t,Context})$, $mlogit(P_{t,EV})$ or $mlogit(P_{t,EV_{back}})$ to Eq. 3, would help to explain the behavioral accuracy better. Because the analysis was split for congruent and incongruent trials, we excluded the terms involving a Congruency effect. For incongruent trials, only $logit(P_{t,Context})$ improved the fit (LR-tests: $logit(P_{t,Context})$: $\chi^2_{(1)} = 3.66$, $p = .055$, $mlogit(P_{t,EV})$: $\chi^2_{(1)} = 0.28$, $p = .599$, $mlogit(P_{t,EV_{back}})$: $\chi^2_{(1)} = 0.0$, $p = .957$). In a second step we then separately tested the interactions $logit(P_{t,Context}) \times mlogit(P_{t,EV})$ or $logit(P_{t,Context}) \times mlogit(P_{t,EV_{back}})$ and found that only the latter had improved the fit ($\chi^2_{(1)} = 1.78$, $p = .183$, $\chi^2_{(1)} = 6.33$, $p = .012$, respectively). For congruent trials, only $mlogit(P_{t,EV_{back}})$ and marginally $mlogit(P_{t,EV})$ improved the fit (LR-tests: $logit(P_{t,Context})$: $\chi^2_{(1)} = 0.0$, $p = .922$, $mlogit(P_{t,EV})$: $\chi^2_{(1)} = 3.5$, $p = .061$, $mlogit(P_{t,EV_{back}})$: $\chi^2_{(1)} = 6.48$, $p = .011$). In a second step we tested separately the interactions $logit(P_{t,Context}) \times mlogit(P_{t,EV})$,$logit(P_{t,Context}) \times mlogit(P_{t,EV_{back}})$ or $mlogit(P_{t,EV_{back}}) \times mlogit(P_{t,EV})$ and found none of these improved model fit when adding them to a model that included both main effects from the previous step ($\chi^2_{(1)} = 0.34$, $p = .560$, $\chi^2_{(1)} = .278$, $p = .598$, $\chi^2_{(1)} = 2.49$, $p = .115$, respectively).

References

[1] Daniel Kahneman and Amos Tversky. Prospect theory : An analysis of decisions under risk. *Econometrica*, 47:278, 1979.

[2] John O'Doherty, Morten L Kringelbach, Edmund T Rolls, Julia Hornak, and Caroline Andrews. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience*, 4(1):95–102, 2001.

[3] Camillo Padoa-Schioppa and John A Assad. Neurons in the orbitofrontal cortex encode economic value. *Nature*, 441 (7090):223–226, 2006.

[4] Oscar Bartra, Joseph T McGuire, and Joseph W Kable. The valuation system: a coordinate-based meta-analysis of bold fmri experiments examining neural correlates of subjective value. *Neuroimage*, 76:412–427, 2013. ISSN 1053-8119.

[5] John A Clithero and Antonio Rangel. Informatic parcellation of the network involved in the computation of subjective value. *Social cognitive and affective neuroscience*, 9(9):1289–1302, 2014.

[6] Hilke Plassmann, John O'doherty, and Antonio Rangel. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of neuroscience*, 27(37):9984–9988, 2007.

[7] Erin L Rich and Jonathan D Wallis. Decoding subjective decisions from orbitofrontal cortex. *Nature neuroscience*, 19 (7):973–980, 2016.

[8] Sebastien Ballesta, Weikang Shi, Katherine E Conen, and Camillo Padoa-Schioppa. Values encoded in orbitofrontal cortex are causally related to economic choices. *Nature*, 2020.

[9] Maurizio Corbetta and Gordon L Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3):201–215, 2002.

[10] Mark G Stokes, Makoto Kusunoki, Natasha Sigala, Hamed Nili, David Gaffan, and John Duncan. Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78(2):364–375, 2013.

[11] Yael Niv, Reka Daniel, Andra Geana, Samuel J Gershman, Yuan Chang Leong, Angela Radulescu, and Robert C Wilson. Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, 35 (21):8145–8157, 2015. ISSN 0270-6474.

[12] Yuan Chang Leong, Angela Radulescu, Reka Daniel, Vivian DeWoskin, and Yael Niv. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, 93(2):451–463, 2017.

[13] Peter H Rudebeck and Elisabeth A Murray. The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. *Neuron*, 84(6):1143–1156, 2014.

[14] Vikram S. Chib, Antonio Rangel, Shinsuke Shimojo, and John P. O'Doherty. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39): 12315–12320, 2009. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.2575-09.2009.

[15] Daniel McNamee, Antonio Rangel, and John P O'doherty. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature neuroscience*, 16(4):479–485, 2013. ISSN 1097-62656.

[16] Gabriel Pelletier and Lesley K Fellows. A critical role for human ventromedial frontal lobe in value comparison of complex objects based on attribute configuration. *Journal of Neuroscience*, 39(21):4124–4132, 2019.

[17] Thorsten Kahnt, Jakob Heinzle, Soyoung Q Park, and John-Dylan Haynes. Decoding different roles for vmpfc and dlpfc in multi-attribute decision making. *Neuroimage*, 56(2):709–715, 2011.

[18] Ulrike Basten, Guido Biele, Hauke R. Heekeren, and Christian J. Fiebach. How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, 107(50):21767–21772, 2010. doi: 10.1073/pnas.0908104107.

[19] Nitzan Shahar, Rani Moran, Tobias U Hauser, Rogier A Kievit, Daniel McNamee, Michael Moutoussis, Raymond J Dolan, NSPN Consortium, et al. Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proceedings of the National Academy of Sciences*, 116(32):15871–15876, 2019.

[20] Vickie Li, Elizabeth Michael, Jan Balaguer, Santiago Herce Castañón, and Christopher Summerfield. Gain control explains the effect of distraction in human perceptual, cognitive, and economic decision making. *Proceedings of the National Academy of Sciences*, 115(38):E8825–E8834, 2018. ISSN 0027-8424.

[21] Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78, 2013. ISSN 1476-4687.

[22] Yu Takagi, Laurence T Hunt, Mark W Woolrich, Timothy EJ Behrens, and Miriam Klein-Flugge. Projections of non-invasive human recordings into state space show unfolding of spontaneous and over-trained choice. *bioRxiv*, 2020.

[23] Nicolas W Schuck, Robert Gaschler, Dorit Wenke, Jakob Heinzle, Peter A Frensch, John-Dylan Haynes, and Carlo Reverberi. Medial prefrontal cortex predicts internally driven strategy shifts. *Neuron*, 86(1):331–340, 2015.

[24] Brian A Anderson. A value-driven mechanism of attentional selection. *Journal of vision*, 13(3):7–7, 2013.

[25] Marcus Grueschow, Rafael Polania, Todd A Hare, and Christian C Ruff. Automatic versus choice-dependent value representations in the human brain. *Neuron*, 85(4):874–885, 2015.

[26] Maël Lebreton, Soledad Jorge, Vincent Michel, Bertrand Thirion, and Mathias Pessiglione. An automatic valuation system in the human brain: Evidence from functional neuroimaging. *Neuron*, 64(3):431 – 439, 2009. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2009.09.040.

[27] Margaret L Schlichting, Jeanette A Mumford, and Alison R Preston. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature communications*, 6 (1):1–10, 2015.

[28] Geoffrey Schoenbaum and Matthew Roesch. Orbitofrontal cortex, associative learning, and expectancies. *Neuron*, 47 (5):633–636, 2005.

[29] Michael L Mack, Alison R Preston, and Bradley C Love. Ventromedial prefrontal cortex compression during concept learning. *Nature communications*, 11(1):1–11, 2020.

[30] Christian F Doeller, Caswell Barry, and Neil Burgess. Evidence for grid cells in a human memory network. *Nature*, 463 (7281):657–661, 2010.

[31] Alexandra O Constantinescu, Jill X O'Reilly, and Timothy EJ Behrens. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292):1464–1468, 2016.

[32] Nicolas W Schuck, Ming Bo Cai, Robert C Wilson, and Yael Niv. Human orbitofrontal cortex represents a cognitive map of state space. *Neuron*, 91(6):1402–1412, 2016.

[33] Stephanie CY Chan, Yael Niv, and Kenneth A Norman. A probability distribution over latent causes, in the orbitofrontal cortex. *Journal of Neuroscience*, 36(30):7817–7828, 2016.

[34] Nicolas W Schuck, Robert Wilson, and Yael Niv. A state representation for reinforcement learning and decision-making in the orbitofrontal cortex. In *Goal-directed decision making*, pages 259–278. Elsevier, 2018.

[35] G Elliott Wimmer and Christian Büchel. Learning of distant state predictions by the orbitofrontal cortex in humans. *Nature communications*, 10(1):1–11, 2019.

[36] Jingfeng Zhou, Matthew PH Gardner, Thomas A Stalnaker, Seth J Ramus, Andrew M Wikenheiser, Yael Niv, and Geoffrey Schoenbaum. Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. *Current Biology*, 29(6):897–907, 2019.

[37] Anja Farovik, Ryan J Place, Sam McKenzie, Blake Porter, Catherine E Munro, and Howard Eichenbaum. Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. *Journal of Neuroscience*, 35(21):8333–8344, 2015.

[38] Praveen K Pilly and Aaron R Seitz. What a difference a parameter makes: A psychophysical comparison of random dot motion algorithms. *Vision research*, 49(13):1599–1612, 2009. ISSN 0042-6989.

[39] Denis Cousineau et al. Confidence intervals in within-subject designs: A simpler solution to loftus and masson's method. *Tutorials in quantitative methods for psychology*, 1(1):42–45, 2005.

[40] Richard D Morey et al. Confidence intervals from normalized data: A correction to cousineau (2005). *reason*, 4(2): 61–64, 2008.

[41] Laurence T Hunt, Nils Kolling, Alireza Soltani, Mark W Woolrich, Matthew FS Rushworth, and Timothy EJ Behrens. Mechanisms underlying cortical activity during value-guided choice. *Nature neuroscience*, 15(3):470–476, 2012. ISSN 1097-6256.

[42] Satohiro Tajima, Jan Drugowitsch, and Alexandre Pouget. Optimal policy for value-based decision-making. *Nature Communications*, 7:12400, 2016. doi: 10.1038/ncomms12400 https://www.nature.com/articles/ncomms12400supplementary-information. URL `https://doi.org/10.1038/ncomms12400`.

[43] Ian Krajbich, Björn Bartling, Todd Hare, and Ernst Fehr. Rethinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6:7455, 2015. doi: 10.1038/ncomms8455. URL `https://doi.org/10.1038/ncomms8455`.

[44] Arni Magnusson, Hans Skaug, Anders Nielsen, Casper Berg, Kasper Kristensen, Martin Maechler, Koen van Bentham, Ben Bolker, Mollie Brooks, and Maintainer Mollie Brooks. Package 'glmmtmb'. *R Package Version 0.2. 0*, 2017.

[45] Robert C Wilson, Yuji K Takahashi, Geoffrey Schoenbaum, and Yael Niv. Orbitofrontal cortex as a cognitive map of task space. *Neuron*, 81(2):267–279, 2014.

[46] Yael Niv. Learning task-state representations. *Nature neuroscience*, 22(10):1544–1553, 2019.

[47] Arno Klein, Satrajit S. Ghosh, Forrest S. Bao, Joachim Giard, Yrjö Häme, Eliezer Stavsky, Noah Lee, Brian Rossa, Martin Reuter, Elias Chaibub Neto, and Anisha Keshavan. Mindboggling morphometry of human brains. *PLOS Computational Biology*, 13(2):e1005350, 2017. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1005350. URL `http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005350`.

[48] Anders M. Dale, Bruce Fischl, and Martin I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999. ISSN 1053-8119. doi: 10.1006/nimg.1998.0395. URL `http://www.sciencedirect.com/science/article/pii/S1053811998903950`.

[49] David H Brainard and Spatial Vision. The psychophysics toolbox. *Spatial vision*, 10:433–436, 1997.

[50] Denis G Pelli and Spatial Vision. The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10:437–442, 1997.

[51] Mario Kleiner, David Brainard, Denis Pelli, Allen Ingling, Richard Murray, Christopher Broussard, et al. What's new in psychtoolbox-3. *Perception*, 36(14):1, 2007.

[52] *MATLAB version 9.3.0.713579 (R2017b)*. The Mathworks, Inc., Natick, Massachusetts, 2017.

[53] *MATLAB version (R2012b)*. The Mathworks, Inc., Natick, Massachusetts, 2017.

[54] Joshua T Abbott, Thomas L Griffiths, and Terry Regier. Focal colors across languages are representative members of color categories. *Proceedings of the National Academy of Sciences*, 113(40):11178–11183, 2016. ISSN 0027-8424.

[55] Helen C Barron, Mona M Garvert, and Timothy EJ Behrens. Repetition suppression: a means to index neural representations using bold? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1705): 20150355, 2016.

[56] Mona M Garvert, Raymond J Dolan, and Timothy EJ Behrens. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife*, 6:e17086, 2017.

[57] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL `https://www.R-project.org/`.

[58] RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, PBC., Boston, MA, 2020. URL `http://www.rstudio.com/`.

[59] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015. doi: 10.18637/jss.v067.i01.

[60] Nikolaus Weiskopf, Chloe Hutton, Oliver Josephs, and Ralf Deichmann. Optimal epi parameters for reduction of susceptibility-induced bold sensitivity losses: A whole-brain analysis at 3 t and 1.5 t. *NeuroImage*, 33(2):493 – 504, 2006. ISSN 1053-8119. doi: https://doi.org/10.1016/j.neuroimage.2006.07.029.

[61] Krzysztof J Gorgolewski, Tibor Auer, Vince D Calhoun, R Cameron Craddock, Samir Das, Eugene P Duff, Guillaume Flandin, Satrajit S Ghosh, Tristan Glatard, Yaroslav O Halchenko, et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific data*, 3(1):1–9, 2016.

[62] Xiangrui Li, Paul S Morgan, John Ashburner, Jolinda Smith, and Christopher Rorden. The first step for neuroimaging data analysis: Dicom to nifti conversion. *Journal of neuroscience methods*, 264:47–56, 2016.

[63] Oscar Esteban, Daniel Birman, Marie Schaer, Oluwasanmi O Koyejo, Russell A Poldrack, and Krzysztof J Gorgolewski. Mriqc: Advancing the automatic prediction of image quality in mri from unseen sites. *PloS one*, 12(9):e0184661, 2017.

[64] Oscar Esteban, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, Mathias Goncalves, Elizabeth DuPre, Madeleine Snyder, Hiroyuki Oya, Satrajit Ghosh, Jessey Wright, Joke Durnez, Russell Poldrack, and Krzysztof Jacek Gorgolewski. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 2018. doi: 10.1038/s41592-018-0235-4.

[65] Oscar Esteban, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, Asier Erramuzpe, Mathias Kent, James D. andGoncalves, Elizabeth DuPre, Kevin R. Sitek, Daniel E. P. Gomez, Daniel J. Lurie, Zhifang Ye, Russell A. Poldrack, and Krzysztof J. Gorgolewski. fmriprep. *Software*, 2018. doi: 10.5281/zenodo.852659.

[66] K. Gorgolewski, C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5:13, 2011. doi: 10.3389/fninf.2011.00013.

[67] Krzysztof J. Gorgolewski, Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, Hans Johnson, Christopher Burns, Alexandre Manhães-Savio, Carlo Hamalainen, Benjamin Yvernault, Taylor Salo, Kesshi Jordan, Mathias Goncalves, Michael Waskom, Daniel Clark, Jason Wong, Fred Loney, Marc Modat, Blake E Dewey, Cindee Madison, Matteo Visconti di Oleggio Castello, Michael G. Clark, Michael Dayan, Dav Clark, Anisha Keshavan, Basile Pinsard, Alexandre Gramfort, Shoshana Berleant, Dylan M. Nielson, Salma Bougacha, Gael Varoquaux, Ben Cipollini, Ross Markello, Ariel Rokem, Brendan Moloney, Yaroslav O. Halchenko, Demian Wassermann, Michael Hanke, Christian Horea, Jakub Kaczmarzyk, Gilles de Hollander, Elizabeth DuPre, Ashley Gillman, David Mordom, Colin Buchanan, Rosalia Tungaraza, Wolfgang M. Pauli, Shariq Iqbal, Sharad Sikka, Matteo Mancini, Yannick Schwartz, Ian B. Malone, Mathieu Dubois, Caroline Frohlich, David Welch, Jessica Forbes, James Kent, Aimi Watanabe, Chad Cumba, Julia M. Huntenburg, Erik Kastman, B. Nolan Nichols, Arman Eshaghi, Daniel Ginsburg, Alexander Schaefer, Benjamin Acland, Steven Giavasis, Jens Kleesiek, Drew Erickson, René Küttner, Christian Haselgrove, Carlos Correa, Ali Ghayoor, Franz Liem, Jarrod Millman, Daniel Haehn, Jeff Lai, Dale Zhou, Ross Blair, Tristan Glatard, Mandy Renfro, Siqi Liu, Ari E. Kahn, Fernando Pérez-García, William Triplett, Leonie Lampe, Jörg Stadler, Xiang-Zhen Kong, Michael Hallquist, Andrey Chetverikov, John Salvatore, Anne Park, Russell Poldrack, R. Cameron Craddock, Souheil Inati, Oliver Hinds, Gavin Cooper, L. Nathan Perkins, Ana Marina, Aaron Mattfeld, Maxime Noel, Lukas Snoek, K Matsubara, Brian Cheung, Simon Rothmei, Sebastian Urchs, Joke Durnez, Fred Mertz, Daniel Geisler, Andrew Floren, Stephan Gerhard, Paul Sharp, Miguel Molina-Romero, Alejandro

Weinstein, William Broderick, Victor Saase, Sami Kristian Andberg, Robbert Harms, Kai Schlamp, Jaime Arias, Dimitri Papadopoulos Orfanos, Claire Tarbert, Arielle Tambini, Alejandro De La Vega, Thomas Nickson, Matthew Brett, Marcel Falkiewicz, Kornelius Podranski, Janosch Linkersdörfer, Guillaume Flandin, Eduard Ort, Dmitry Shachnev, Daniel McNamee, Andrew Davison, Jan Varada, Isaac Schwabacher, John Pellman, Martin Perez-Guevara, Ranjeet Khanuja, Nicolas Pannetier, Conor McDermottroe, and Satrajit Ghosh. Nipype. *Software*, 2018. doi: 10.5281/zenodo.596855.

[68] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8, 2014. ISSN 1662-5196. doi: 10.3389/fninf.2014.00014. URL `https://www.frontiersin.org/articles/10.3389/fninf.2014.00014/full`.

[69] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4itk: Improved n3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010. ISSN 0278-0062. doi: 10.1109/TMI.2010.2046908.

[70] VS Fonov, AC Evans, RC McKinstry, CR Almli, and DL Collins. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, Supplement 1:S102, 2009. ISSN 1053-8119. doi: 10.1016/S1053-8119(09)70884-5. URL `http://www.sciencedirect.com/science/article/pii/S1053811909708845`.

[71] B.B. Avants, C.L. Epstein, M. Grossman, and J.C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1): 26–41, 2008. ISSN 1361-8415. doi: 10.1016/j.media.2007.06.004. URL `http://www.sciencedirect.com/science/article/pii/S1361841507000606`.

[72] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1):45–57, 2001. ISSN 0278-0062. doi: 10.1109/42.906424.

[73] Robert W. Cox and James S. Hyde. Software tools for analysis and visualization of fmri data. *NMR in Biomedicine*, 10(4-5):171–178, 1997. doi: 10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

[74] Douglas N Greve and Bruce Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, 2009. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2009.06.060.

[75] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–841, 2002. ISSN 1053-8119. doi: 10.1006/nimg.2002.1132. URL `http://www.sciencedirect.com/science/article/pii/S1053811902911328`.

[76] Jonathan D. Power, Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. Methods to detect, characterize, and remove motion artifact in resting state fmri. *NeuroImage*, 84(Supplement C):320–341, 2014. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2013.08.048. URL `http://www.sciencedirect.com/science/article/pii/S1053811913009117`.

[77] Yashar Behzadi, Khaled Restom, Joy Liau, and Thomas T. Liu. A component based noise correction method (CompCor) for BOLD and perfusion based fmri. *NeuroImage*, 37(1):90–101, 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2007.04.042. URL `http://www.sciencedirect.com/science/article/pii/S1053811907003837`.

[78] C. Lanczos. Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1):76–85, 1964. ISSN 0887-459X. doi: 10.1137/0701007. URL `http://epubs.siam.org/doi/10.1137/0701007`.

[79] Gary H Glover, Tie-Qiang Li, and David Ress. Image-based method for retrospective correction of physiological motion effects in fmri: Retroicor. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 44(1):162–167, 2000.

[80] Chloe Hutton, Oliver Josephs, Jörg Stadler, Eric Featherstone, Alphonso Reid, Oliver Speck, Johannes Bernarding, and Nikolaus Weiskopf. The impact of physiological noise correction on fmri at 7 t. *Neuroimage*, 57(1):101–112, 2011.

[81] Ann K Harvey, Kyle TS Pattinson, Jonathan CW Brooks, Stephen D Mayhew, Mark Jenkinson, and Richard G Wise. Brainstem functional magnetic resonance imaging: disentangling signal from physiological noise. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 28(6): 1337–1344, 2008.

[82] Lars Kasper, Steffen Bollmann, Andreea O Diaconescu, Chloe Hutton, Jakob Heinzle, Sandra Iglesias, Tobias U Hauser, Miriam Sebold, Zina-Mary Manjaly, Klaas P Pruessmann, et al. The physio toolbox for modeling physiological noise in fmri data. *Journal of neuroscience methods*, 276:56–72, 2017.

[83] William D Penny, Karl J Friston, John T Ashburner, Stefan J Kiebel, and Thomas E Nichols. *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011.

[84] Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Olivier Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

[85] Edmund T Rolls, Marc Joliot, and Nathalie Tzourio-Mazoyer. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *Neuroimage*, 122:1–5, 2015.

[86] Edmund T Rolls, Chu-Chung Huang, Ching-Po Lin, Jianfeng Feng, and Marc Joliot. Automated anatomical labelling atlas 3. *NeuroImage*, 206:116189, 2020.

[87] Jeanette A Mumford, Jean-Baptiste Poline, and Russell A Poldrack. Orthogonalization of regressors in fmri models. *PloS one*, 10(4):e0126255, 2015.

[88] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[89] Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler, and Benjamin M. Bolker. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2):378–400, 2017. URL https://journal.r-project.org/archive/2017/RJ-2017-066/index.html.

Supplementary Information

- Fig. S1: Full procedure and experimental design for all phases, related to Fig 1
- Fig. S2: Nested RT models, related to Fig 2
- Fig. S3: Alternative RT models, extended RT model comparisons and correlation matrix of all regressors, related to Fig 2
- Fig. S4: Exploratory analysis of RT model presented in Main Text, related to Fig 2
- Fig. S5: Behavioral accuracy results: related to Fig 2
- Fig. S6: Supplementary information for Value similarity analysis: related to Fig. 4
- Fig. S7: Supplementary information for perceptual similarity analysis: related to Fig. 4
- Fig. S8: Modelling probability assigned to the EV class: related to Fig. 5
- Fig. S9: Additional univariate results, related to Fig. 8
- Table S1: Detailed univariate results: Clusters for whole brain univariate analysis, related to Fig. 8

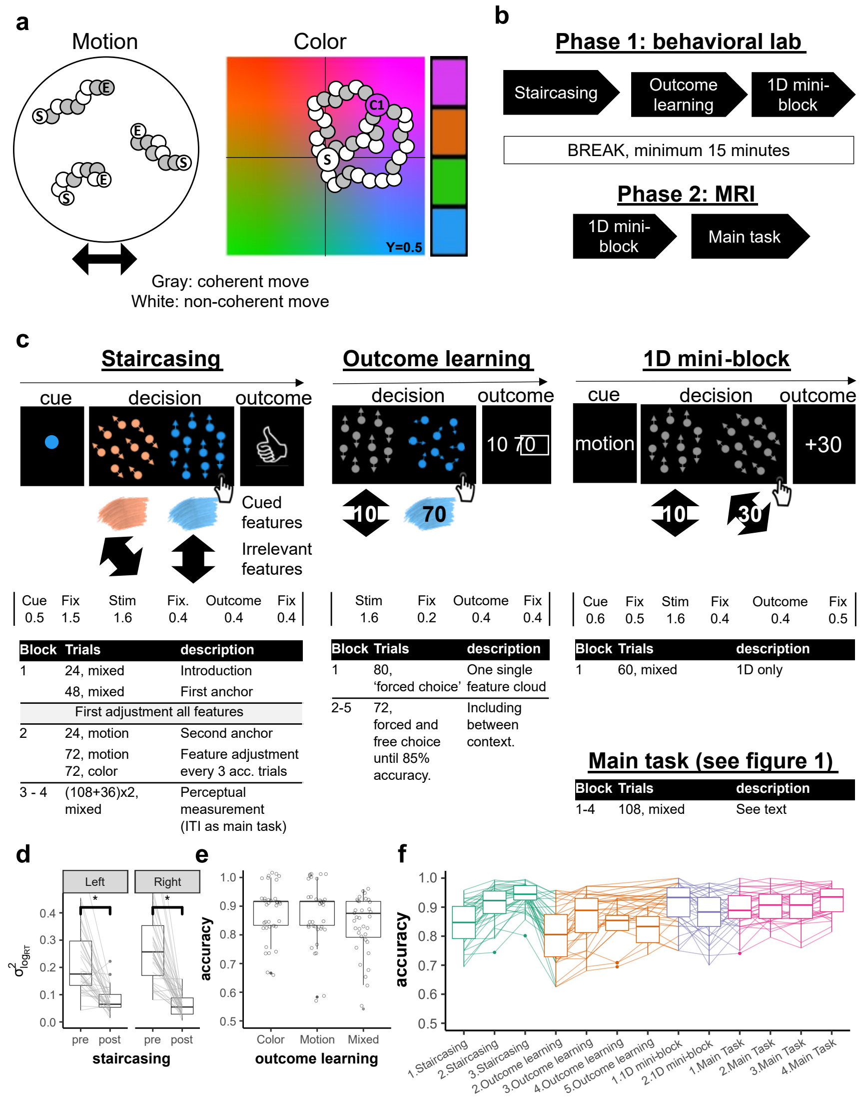


Figure S1: Full procedure and experimental design for all phases

Figure S1: Full procedure and experimental design for all phases, related to Fig 1. a. Brownian algorithm for color and motion. Each illustration shows the course of 3 example dots; 'S' and 'E' marked dots reflect Start and End positions, respectively. Remaining dots represent location in space for different frames. Left panel: Horizontal motion trial. Shown are framewise dot positions between start and end. In each frame, a different set of dots moved coherently in the designated direction (gray) with a fixed speed; remaining dots moved in a random direction [conceptually taken from 38]. Right panel: Example of a pink color trial. We simulated the YCbCr color space that is believed to represent the human perception in a relative accurate way [cf. 54]. A fixed luminance of Y = 0.5 was used. For technical reasons we sliced the X-axis by 0.1 on each side and the Y-axis by 0.2 from the bottom of the space to ensure the middle of the space remained gray given the chosen luminance. In each frame, a different set of dots (always 30% of the dots) moved coherently towards the target color in a certain speed whereas the rest were assigned with a random direction. All target colors were offset by 23.75% from the center towards each corner. Right bar illustrates the used target colors. **b.** Full procedure. The experiment consisted of two phases, the first one took place in the behavioral lab and included Staircasig, Outcome-learning and the first 1D mini-block. The second took place inside the MRI scanner and consisted of the second 1D mini-block and the main task. **c.** Example trial procedures and timing of the different tasks. Timing of each trial is depicted below illustrations. **Staircasing (left)** Each trial started with a cue of the relevant feature. Each cloud had one or two features (motion and/or color) and participants had to detect the cued feature. Participants' task was to choose the cued feature (here: blue). After a choice, participants received feedback if they were correct and faster than 1 second, correct and slower, or wrong. **Outcome learning (middle)** Participants were presented with either one or two single-feature clouds and asked to chose the highest valued feature. Following their choice, they were presented with the values of both clouds, with the chosen cloud's associated value marked with a square around it. The pair of shown stimuli included across contexts comparisons, e.g. between up/right and blue, as shown. **1D mini block (right)** At the end of the first phase and beginning of the second phase participants completed a mini-block of 60 1D trials during the anatomical scan (30 color-only, 30 motion-only, interleaved). Participants were again asked to make a value-based two alternative forced choice choice decision. In each trial, they were first presented with a contextual cue (color/motion), followed by the presentation of two single-feature clouds of the cued context. After a choice, they were presented with the chosen-cloud's value. No BOLD response was measured during these blocks and timing of the trials was fixed and shorter than in the main task (see Main task preparation in online methods) **Main task (bottom)** This part included 4 blocks, each consisting of 36 1D and 72 2D trials trials presented in an interleaved fashion (see online method and Fig. 1). **d.** Button specific reduction in RT variance following the staircasing. We verified that the staircasing procedure also reduced differences in detection speed between features when testing each button separately. Depicted is the variance of reaction times (RTs) across different color and motion features (y axis). While participants' RTs were markedly different for different features before staircasing (pre), a significant reduction in RT differences was observed after the procedure (post, $p < .001$.) **e.** Choice accuracy in outcome learning trials. Participants achieved near ceiling accuracy in choosing the highest valued feature in the outcome learning task, also when testing for color, motion and mixed trials separately (ps$< .001$). Mixed trials only appeared in this part of the experiment to encourage mapping of the values on similar scales. **f.** Accuracy throughout the experiment, plotted for each block of each part of the experiment. *In the staircasing (left)* High accuracy for the adjustment and measurement blocks (2-3) ensured that there were no difficulties in perceptual detection of the features. *In Outcome learning* a clear increase in accuracy throughout this task indicated learning of feature-outcome associations. Note that Block 5 of this part was only included for those who did not achieve 85% accuracy beforehand. Starting the *1D mini blocks* (middle) and throughout the*main task* (right) until the end of the experiment high accuracy. μ and σ from left to right: Staircasing: .84,.07;.91,.06;.94,.04; Outcome Learning: .81,.1;.86,.09;.83,.08;.82,.06; 1D mini blocks: .91,.07;.88,.08; Main task: .89,.06;.91,.05;.9,.06;.92,.05.

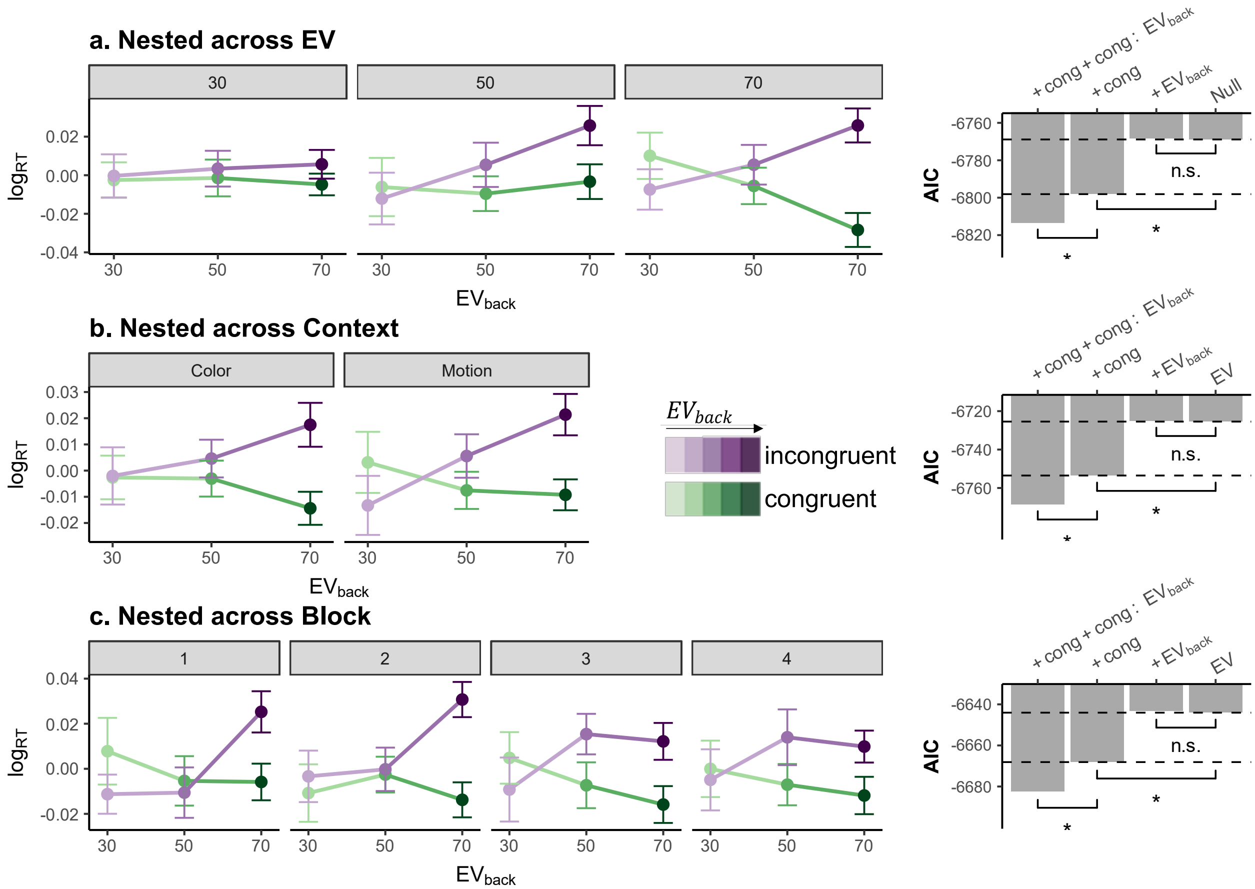


Figure S2: Nested RT models, related to Fig 2

1322 **Figure S2: Nested RT models, related to Fig 2**

1323 **a-c. Nested models within Factors.** Each row represents one congruency analysis, done separately for each
1324 level of expected value (top row), context (middle) or block (bottom). The RT effect of Congruency$_t$ $\times$ EV$_{\mathrm{back}t}$ is
1325 shown on the left, corresponding AICs for mixed effect models with nested factors are shown on the right. RT data
1326 is demeaned for each panel for visual comparison; error bars represent corrected within subject SEMs [39, 40]. Null
1327 models shown on the right are identical to Eq. 2, albeit included $\zeta_{0_{kv}}$, which is the factor-specific (v) intercept
1328 nested within each within each subject level (see online methods). Likelihood ratio tests were performed to asses
1329 improved model fit when adding (1) Congruency or (2) EV$_{\mathrm{back}}$ terms to the Null model and when adding (3)
1330 Congruency $\times$ EV$_{\mathrm{back}}$) in addition to Congruency. Stars represent p values less than .05. For nested within EV, the
1331 Null model did not include a main effect for EV and the LR test was: (1) $\chi^2_{(1)} = 31.22$, $p < .001$; (2) $\chi^2_{(1)} = 1.47$,
1332 $p = .226$; (3) $\chi^2_{(1)} = 19.37$, $p < .001$; For models nested within Context the LR test was: (1) $\chi^2_{(1)} = 30.01$,
1333 $p < .001$; (2) $\chi^2_{(1)} = 1.5$, $p = .22$; (3) $\chi^2_{(1)} = 18.9$, $p < .001$; and for Block: (1) $\chi^2_{(1)} = 26.06$, $p < .001$; (2)
1334 $\chi^2_{(1)} = 1.27$, $p = .26$; (3) $\chi^2_{(1)} = 18.25$, $p < .001$; In the first row (nested across EV) the interaction with EV is
1335 visible, i.e. the higher the EV, the stronger our effects of interests were.

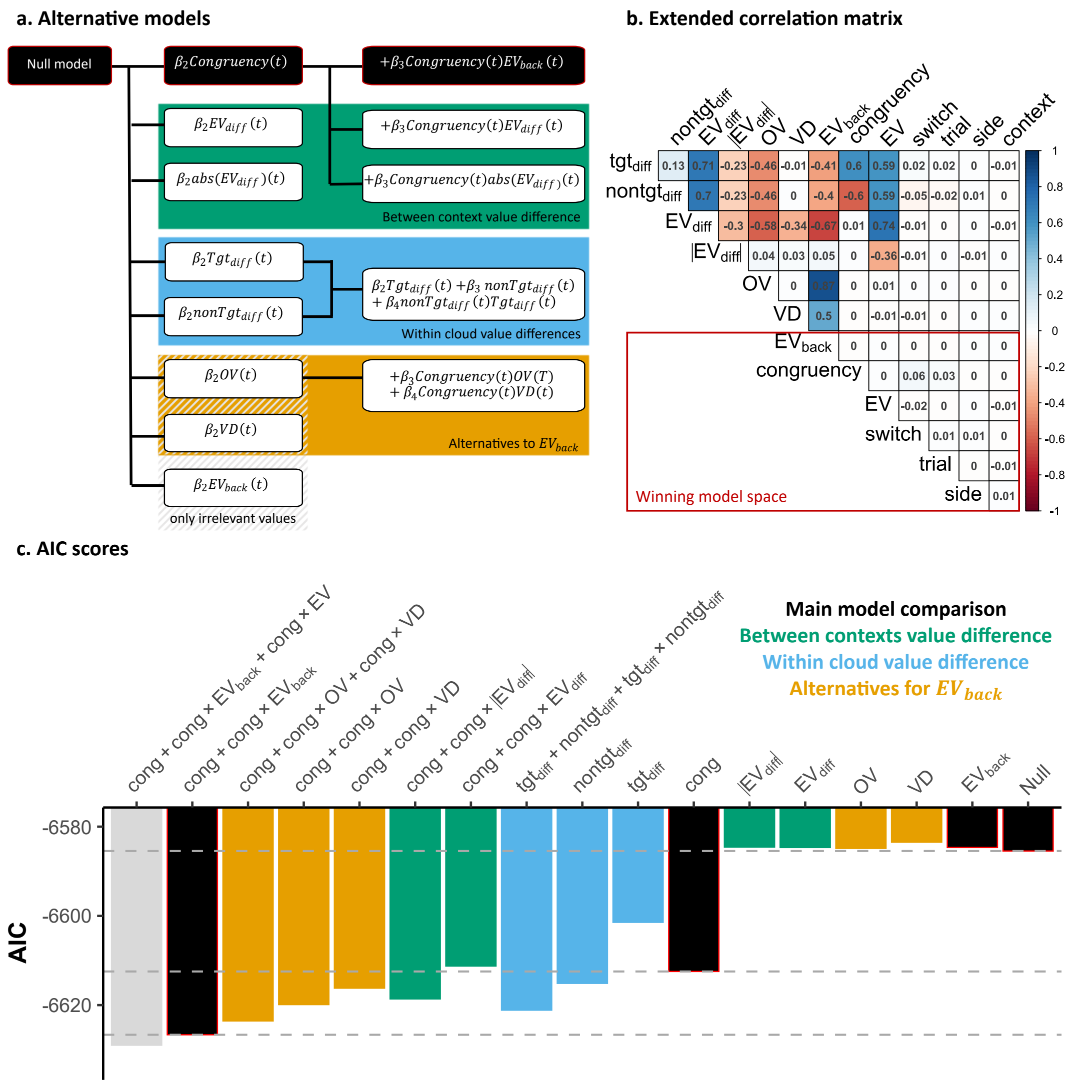


Figure S3: Alternative RT models, extended RT model comparisons and correlation matrix of all regressors, related to Fig 2.

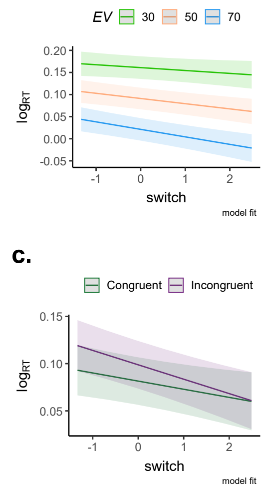
Figure S3: Alternative RT models, extended RT model comparisons and correlation matrix of all regressors, related to Fig 2.
a. Alternative mixed effect models, each represented as a row which lists main factors of interest. We clustered different alternative models into three classes: *Green models* included factors that reflected the difference between the expected values of both contexts (EV - EV_{back}, including unsigned EV factors); *blue* models include instead factor that reflect the value-difference between context within each cloud where 'tgt' (target) is the chosen cloud with the highest value according to the relevant context and *orange* models included two alternative parameterization of values in the non-relevant context: irrelevant features' Value Difference (VD) and Overall Value (OV), which are also orthogonal to Congruency (Cong), and to each other. *In black* is the main model comparison as presented in the main text. **b. Extended correlation matrix**. Averaged correlation across subjects of all scaled regressors for accurate 2D trials (models' input). Marked in red rectangle are main factors of the experiment which are orthogonal by design and used for the model comparison reported in the Main Text. **c. AIC scores.** We tested different alternatives shown in (a) in a stepwise hierarchical model comparison, as in the main text. Each bar represents the AIC (y-axis) of a different model (x-axis) where the labels on the x-axis depict the added terms to the Null model for that specific model. The Null model included nuisance regressors and the main

effect of EV (see ν and β_1 in Eq. 2). The models described in the main text are shown in black. The gray model includes the additional term for Congruency $\times$ EV. Dashed lines correspond to the AIC values of the models used in the main text. Importantly, no main effect representing only the contextually irrelevant values (VD, OV, EV_{back}) nor the difference between the EVs (EV_{diff},$|\text{EV}_{\text{diff}}|$, also when excluding EV from the null model, not presented) improved model fit over the Null model. This supports our finding that neither large irrelevant values, nor their similarity to the objective EV, influenced participants' behavior. Similar to EV_{back}, factors from the green and orange clusters are also orthogonal to Congruency, which allowed us to test their interaction. Factors from the blue cluster highly correlate with both Congruency (and EV_{back}) and therefore were tested separately. Non of the alternatives provided a better AIC score (y axis, lower is better).

a.

effect	Df	AIC	likelihood ratio test	P value
Switch	1	-6490.88	15.19	<.001
Trial	1	-6499.01	7.04	.008
Context	1	-6304.79	201.27	<.001
EV	1	-6159.75	346.31	<.001
Congruency	1	-6496.76	9.29	.002
Congruency x EV	1	-6501.69	4.37	.037
Congruency x EV_{back}	1	-6488.35	17.70	<.001
Switch x Side	1	-6502.42	3.64	.0565
Switch x Context	1	-6503.25	2.80	.094
Switch x EV	1	-6500.23	5.83	.0158
Switch x Congruency	1	-6502.54	3.52	.068
Context x EV	1	-6499.16	31.07	<.001
Context x Trial	1	-6499.16	6.89	.009
Side x EV_{back}	1	-6501.28	4.77	.029

b. **c.**



g. **f.** **e.** **d.**

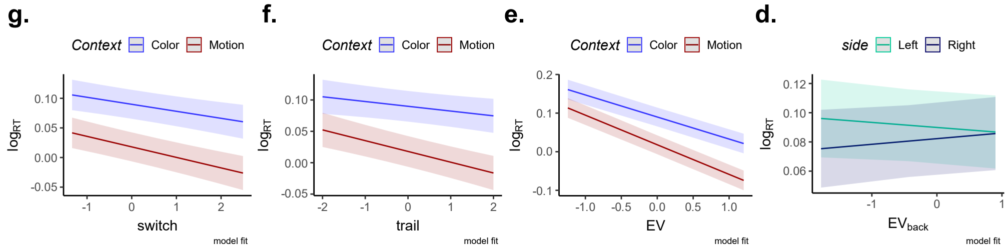
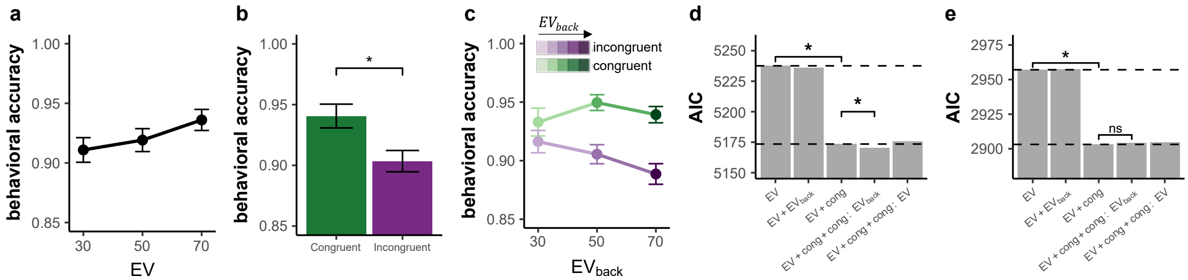


Figure S4: Exploratory analysis of RT model presented in Main Text, related to Fig 2.

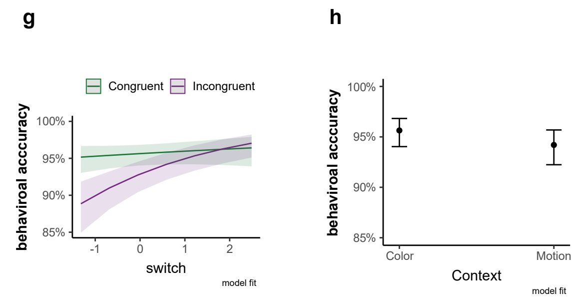
Figure S4: Exploratory analysis of RT model presented in Main Text, related to Fig 2.

a. The table presents the individual contribution of terms taken from Eq. 2 and all possible two-way interactions to the model fit using the drop1 function in R [57]. In short, this exploratory analysis started with a model that included all main effects from Eq. 2 and all possible 2-way interaction between them and tested which terms contribute to the fit. If a term did not improve fit, it was dropped from the model. Presented are all effects with a p value less than $p < .01$. **b-g.** Model fits of all effects with $p < .1$. X-axes are normalized (as in the model) and y-axes reflect RTs on a log scale (model input). Clockwise from the top: RTs became progressively faster with increasing trials since the context switch. This effect was possibly stronger for higher EV (b) and for incongruent trials (c). We note that our experiment was not designed to test the effect of the switch. (d) An interaction of Side and $\mathrm{EV}_{\mathrm{back}}$ was found, for which we offer no explanation. Panels (e) to (g) reflect interaction of context with EV (e), trial (f), and switch (g). We note that due to the used perceptual color space there might be a context-specific ceiling effect in RTs due to training throughout the task which could have induced effects of context. Specifically, since dots start gray and slowly 'gain' the color, it might take a few frames until there is any evidence for color. However, the motion could be theoretically detected already on the second frame (since coherence was very high). This could explain why some effects that represent decrease in RT might hit a boundary for color (and not motion). Crucially, we refer the reader to supplementary Fig S2 where the main model comparison hold also when we ran the model nested within the levels of Context



f

effect	Df	AIC	likelihood ratio test	P value
Context	1	5177.23	4.22	.04
EV	1	5184.14	11,12	.0009
Congruency	1	5186.46	13.44	.0002
Context x Trial	1	5176.76	2.74	.0979
Context x Side	1	5177.01	4.00	.0456
Context x Switch	1	5177.24	4.23	.0398
Context x EV	1	5188.95	15.94	<.0001
Congruency x Switch	1	5185.89	12.87	.0003
Congruency x EV_{back}	1	5176.43	3.41	.0649



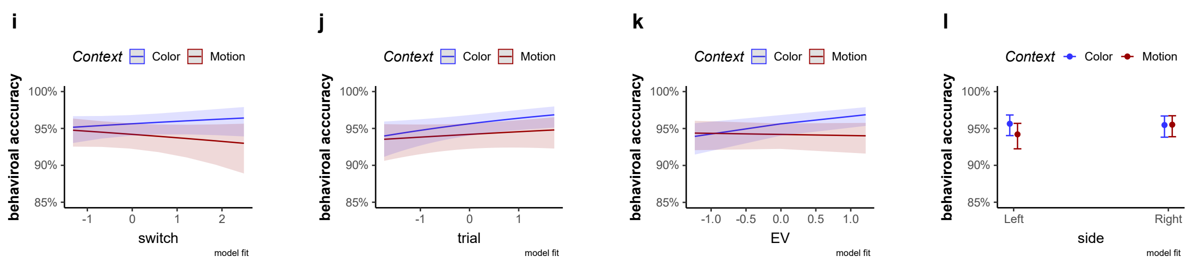
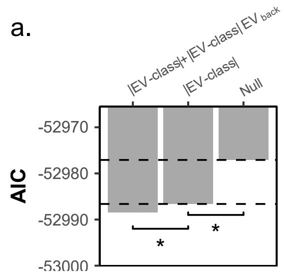


Figure S5: Behavioral accuracy results: related to Fig 2.

Figure S5: Behavioral accuracy results: related to Fig 2.

a. Comparison of accuracy (y-axis) for each level of EV (x-axis) showed that participants were more accurate for higher EV, $p = .001$. **b.** Comparison of congruent versus incongruent trials also revealed a performance benefit of the former, $p = .001$. **c.** The effect of Congruency was modulated by EV_{back}, i.e. the more participants could expect to receive from the ignored context, the less accurate they were when the contexts disagreed (x axis, shades of colours). Further investigations revealed that the modulation of EV_{back} is likely limited to Incongruent trials ($\chi^2_{(1)} = 6.91$, $p = .009$, when modeling only Incongruent trials), yet does not increase accuracy for Congruent trials ($\chi^2_{(1)} = 0.07$, $p = .794$, when modeling only congruent trials), likely due to a ceiling effect. Error bars in panels a-c represent corrected within subject SEMs [39, 40]. **d.** Hierarchical model comparison of choice accuracy, similar to the RT model reported in the main text. These analyses showed that including Congruency improved model fit ($p < .001$). Including the additional interaction of Congruency $\times$ EV_{back} improved the fit even more ($p = .03$). **e.** We replicated the choice accuracy main effect in an independent sample of 21 participants outside of the MRI scanner, i.e. including Congruency improved model fit ($\chi^2_{(1)} = 55.95$, $p < .001$). We did not find a main effect of EV on accuracy in this sample ($\chi^2_{(1)} = 0.93$, $p = .333$). The interaction term Congruency $\times$ EV_{back} did not significantly improve fit in this sample. Modeling only Incongruent trials, as above, reveled that EV_{back} had a marginal effect on accuracy ($\chi^2_{(1)} = 2.90$, $p = .088$). Near-ceiling accuracies in Congruent trials in combination with a smaller sample might have masked the effects. **f.** The table presents the individual contribution of terms taken from Eq. 3 and all possible two-way interactions to the model fit using the drop1 function in R [57]. In short, this exploratory analysis started with a model that included all main effects from Eq. 3 and all possible 2-way interaction between them and tested which terms contribute to the fit. If a term did not improve fit, it was dropped from the model. Subsequent panels present all the effects corresponding to $p < .01$. Note that this is a non-hypothesis driven exploration of the data and that accuracy was very high in general throughout the main task. **g.** Accuracy as a function of time since switch. Akin to RTs, accuracy increased with number of trials since the last context switch, mainly for incongruent trials. **h.** Context effect on accuracy. According to the exploratory model, participants were slightly more accurate in color than in motion trials. However, a direct paired t test between average accuracy of color compared to motion was not significant ($t_{(34)} = 0.96$, $p = .345$) **i-l.** Depicted are some minor interactions of no interest with Context, according to the exploratory model.



b.

Model term	χ^2	Df	Type II Wald χ^2
|EV-class|	11.55	1	<.001
|EV-class| × EV_{back}	3.90	1	.048
|EV-class| × Block	.06	3	.164
|EV-class| × Context	1.45	1	.906
|EV-class| × Switch	1.95	1	.158
|EV-class| × Side	.94	1	.911

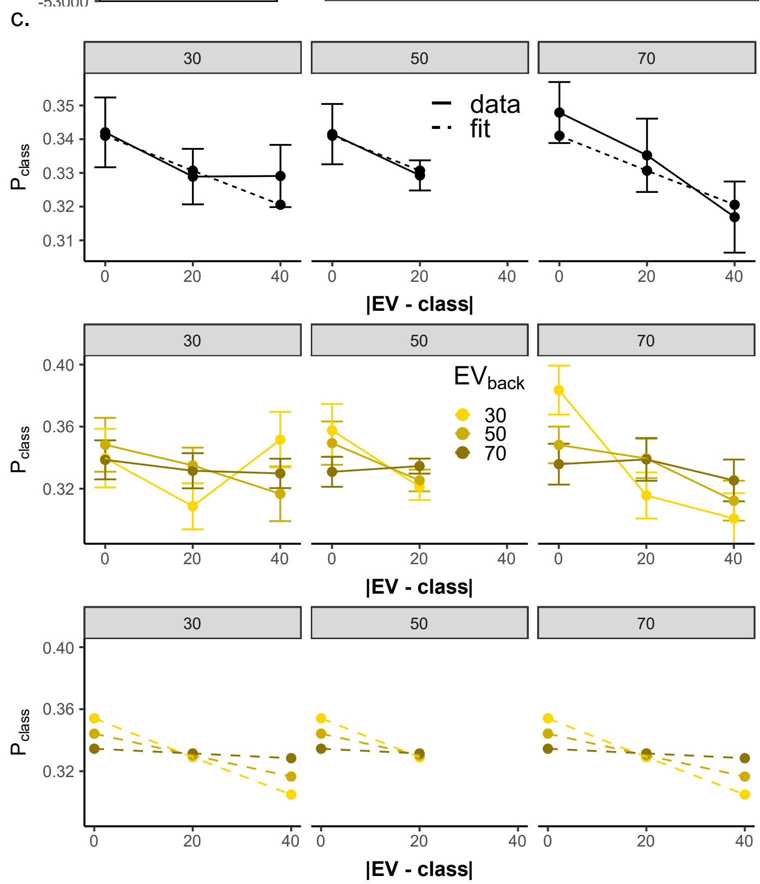
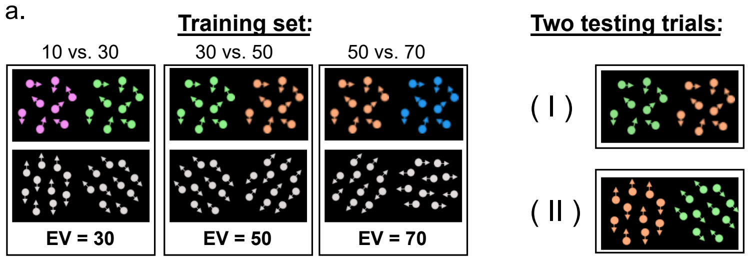


Figure S6: Supplementary information for value similarity analysis: related to Fig. 4

Fig. S6: Supplementary information for Value similarity analysis: related to Fig. 4. **a.** Main value similarity model comparison replicated when fitting the models to unaveraged data. Adding a term for |EV-class| improved model fit (LR test with added term: $\chi^2_{(1)} = 11.56$, $p < .001$). Adding an additional term for |EV-class| × EV_{back} further improved the fit ($\chi^2_{(1)} = 3.86$, $p = .049$), as in the model reported in the main text (Fig. 4b). **b.** Effect of Nuisance regressors on unaveraged data (t, Side, Switch and Context). Same as Congruency and EV_{back}, all of the nuisance regressors don't discriminate between the classes, but rather assign the same value to all three probabilities from that trial (which sum to 1). We therefore tested if any of them modulated the value similarity effect. As can be seen in the table, none of the nuisance regressors modulated the value similarity effect. **c.** Replication of the value similarity model comparison reported in the main text, averaged across nuisance regressors and nested within the levels of EV, i.e. including EV-specific intercepts nested within each within each subject level ($\zeta_{0_{kv}}$, see Online Methods). As in the analysis reported in the Main Text, adding a main effect for |EV-Class| improves model fit ($\chi^2_{(1)} = 16.15$, $p < .001$, first row) as well as adding an additional interaction term |EV-class| × EV_{back} ($\chi^2_{(1)} = 6.16$, $p = .013$, middle row shows data, bottom row shows model fit. Error bars represent corrected within subject SEMs [39, 40]) a.

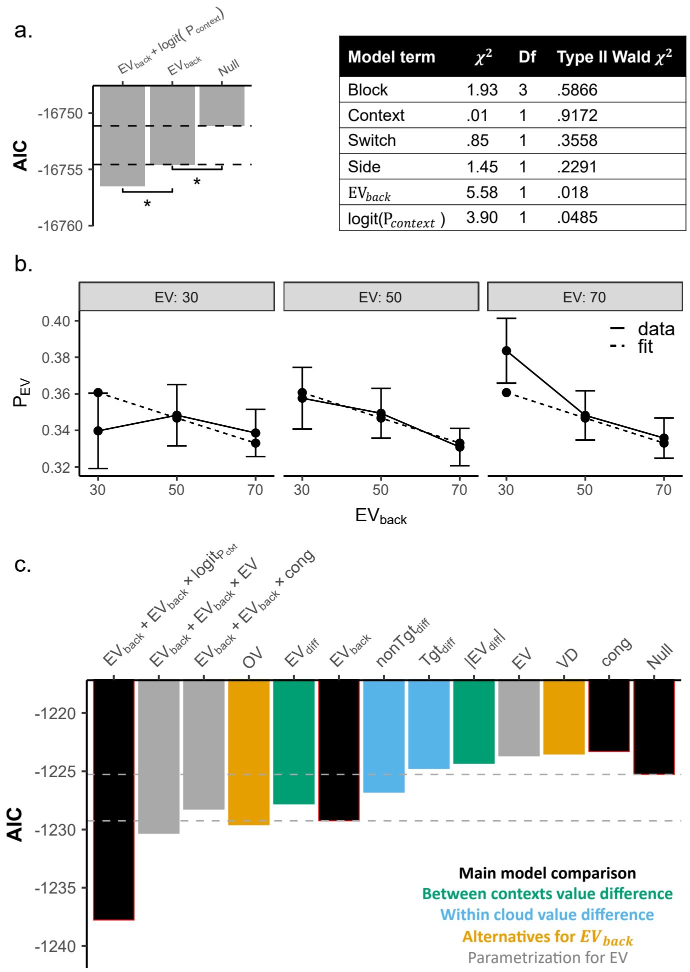


b.

	Test (color)	Test (motion)	*EV*	EV_{back}	Train Class (color & motion)		|EV-class|	Feature Similarity		$Similarity_{back}$
(I) 1D color	30 50		50	--	10	30	**20**	**1**	1x30	-
					30	50	**0**	**2**	1x30, 1x50	-
					50	70	**20**	**1**	1x50	-
(II) 2D color	30 50	10 30	50 color	30 motion	10	30	**20**	**3**	2x30, 1x10	1
					30	50	**0**	**3**	2x30, 1x50	1
					50	70	**20**	**1**	1x50	0

Figure S7: Supplementary information for perceptual similarity analysis: related to Fig. 4

Fig. S7: Supplementary information for perceptual similarity analysis: related to Fig. 4. **a.** Left: training set consisting of 1D trials provided for the classifier for each class (in the experiment the sides were pseudorandomised). Note that each class had the same amount of color and motion 1D trials and that the value difference between the values was always 20. Right: two examples of 2D trials that constituted the classifier test set. **b.** The table illustrates the calculation of feature similarity between classifier test and training in two example trials in one 1D and one 2D trial. Specifically, shown are the corresponding values and features for each trial with the predicted values at each class for the parameters value similarity (|EV-class|), feature similarity and $\text{similarity}_{\text{back}}$. Feature similarity encodes the perceptual overlap between the shown test example and the training examples underlying with each value class. The first row shows a case in which the classifier was tested on a 1D green vs. orange color trial (30 vs 50, EV = 50). Considering in this case for instance the predicted probability that EV=30, the table illustrates the training example underlying the EV = 30 cases (10 vs 30, dark gray shading), the |EV-class| (here: 20, because 50-30), and the feature similarity i.e. how many features from the training class appeared in the test example (here: 1). The second row shows a 2D color trial, reflecting the same value based choice between 30 and 50. The value similarity between training and test stays the same as for the 1D trial shown above. However, the feature similarity between test and training changes because of the motion features. If we take class 30 for example (which is 10 vs 30, dark gray shading), the feature 30 appeared twice (color and motion) and the feature 10 appeared once (motion), i.e. feature similarity now takes on the value 3. $\text{Similarity}_{\text{back}}$ was used to test a perceptual-based alternative to the EV_{back} parameter. $\text{Similarity}_{\text{back}}$ takes on 1 if the perceptual feature corresponding to the EV_{back} appeared in the training class and 0 otherwise (red text in table). As described in the main text, none of the perceptual-similarity encoding alternatives provided a better fit than the reported models that focused on the values the features represent.



Model term	χ^2	Df	Type II Wald χ^2
Block	1.93	3	.5866
Context	.01	1	.9172
Switch	.85	1	.3558
Side	1.45	1	.2291
EV_{back}	5.58	1	.018
logit($P_{context}$)	3.90	1	.0485

Figure S8: Modelling probability assigned to the EV class.

Fig. S8: Modelling probability assigned to the EV class: related to Fig. 5.
a. We replicated the main results using the unaveraged data. The Null model was: $P^k_{t,EV} = \beta_0 + \gamma_{0k} + \nu_1 side(t) + \nu_2 switch(t) + \nu_3 context(t)$, where $P^k_{t,EV}$ is the probability assigned to the class corresponding to the EV of trial t for subject k, β_0 and γ_{0k} represent global and subject-specific intercepts. Side, Switch and Context are the same as in the RT model (Eq. 2); None of these variables had a main effect, $p > 0.4$ (see table, right). The factor $trial$ could not be included due to model convergence issues. Adding a term representing EV_{back} improved model fit (LR test including term: $\chi^2_{(1)} = 5.42$, $p = .019$). Adding an additional term for context decodability further improved the fit ($\chi^2_{(1)} = 3.9$, $p = .048$). The table (right) displays the Type 2 Wald χ^2 test for all main effects from the model. **b.** Depicted is the effect of EV_{back} (x-axis) on the probability assignd to the EV class (P_{EV}, y axis). Solid lines represent the data and dashed lines the model fit of a model that included random effects of subject and EV nested within subject (data averaged across nuisance regressors, adding a main effect for EV_{back} improved model fit ($\chi^2_{(1)} = 5.99$, $p = .014$). Error bars represent corrected within subject SEMs [39, 40]. **c.** Similar to our analysis of alternative models of RT, we clustered models reflecting alternative explanations into three conceptual groups (see color legend; cf. Fig. S3a). All models were fitted to the probability assigned to the objective EV in accurate 2D trials, similar to Eq. 5. Each column represents the AIC (y-axis) of a different model (x-axis) where the labels on the x-axis depict all the main effects included in that specific model (i.e. added to the Null, i.e. Eq. 5 without any main effects). We found no evidence that any other parameters explain the data better than the ones we used in the main text. Specifically, only including main effect of EV_{back}, Overall Value of the irrelevant values (OV) and the difference of both EVs (EV_{diff}) provided a better AIC score than the Null model. Note that adding

OV (-1229.6) only slightly surpassed EV_{back} (-1229.26). Crucially, the correlation of EV_{back} and OV is very high ($\rho = .87$, see main text). We then looked at possible interactions with the EV_{back} effect. Congruency did not seem to modulate the main effect of EV_{back} and adding an interaction term EV $\times$ EV_{back} provided a slightly better AIC (-1230.33), yet this effect was not significant (LR test: $\chi^2_{(1)} = 3.08$, $p = .079$). Section (b) also visualizes this effect. Lastly, adding a term for the Context decodability provided the lowest (i.e. best) AIC score.

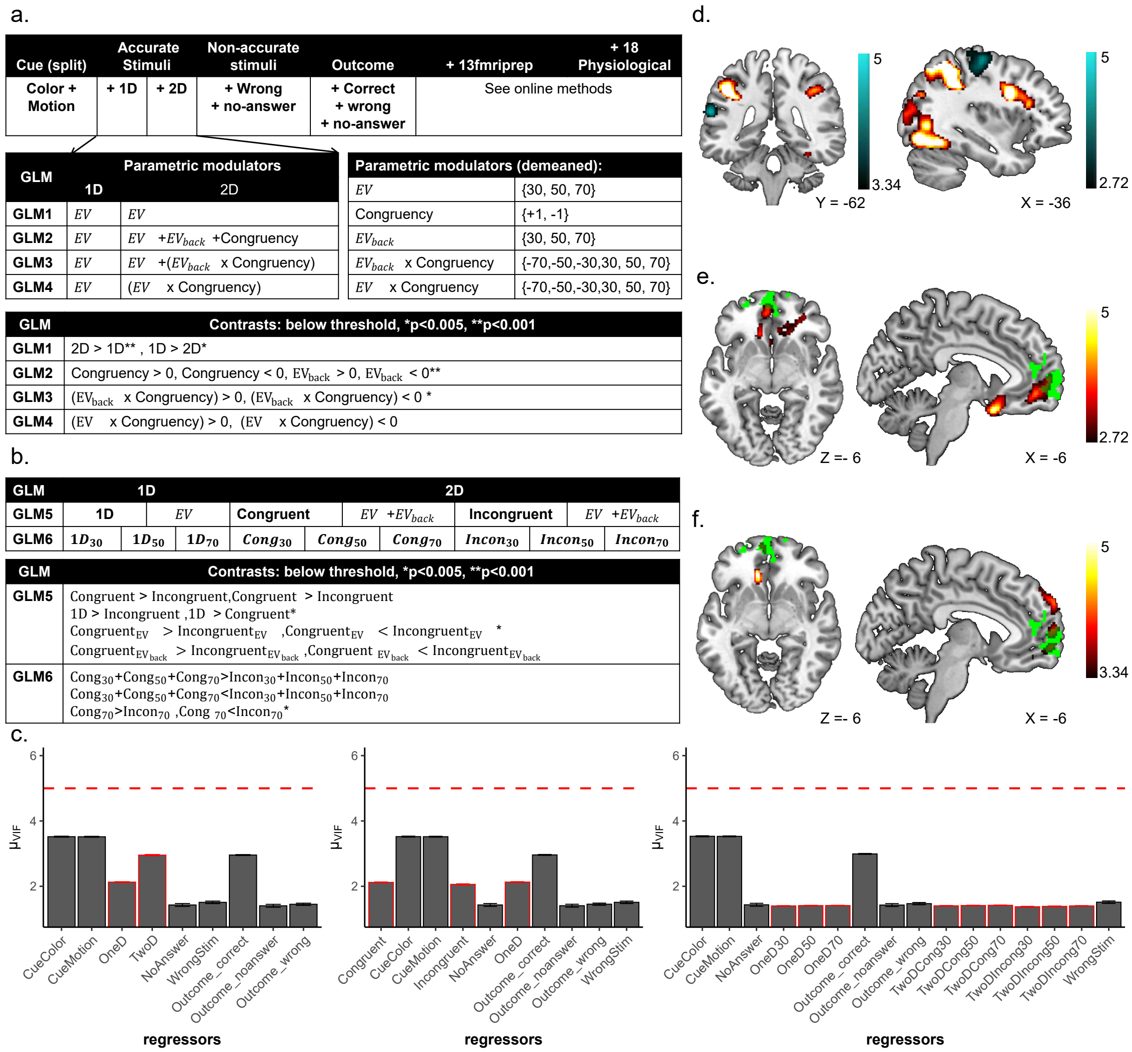


Figure S9: Additional univariate results.

Fig. S9: Additional univariate results, related to Fig. 8.
a. Visualization of GLMs described in the main text. The tables depict the structure of GLMs1-4 which were mainly motivated by the behavioral analysis; onset regressors are shown in the top table, parametric modulators assigned to 1D and 2D onsets (middle-left), the values they were modeled with (demeaned, middle-right) are shown below. The contrasts of interest are shown in the bottom table. The GLMs differed only in their modulations of the 2D trials: GLM1 included only modulators of the objective outcome, GLM2 included one modulator for Congruency and one for EV_{back}, GLM3 included a modulator for the Congruency $\times$ EV_{back} interaction and GLM4 included instead of the EV modulator a modulator of the EV $\times$ Congruency interaction. In the contrast table (bottom) contrasts that only revealed effects at a liberal threshold of $p < .005$ are marked with one star, and contrasts significant at $p < .001$ are marked with **. **b.** We constructed additional GLMs to verify the results of GLMs 1-4. In GLM5 we split the onset of 2D trials into congruent and incongruent trials and assigned a parametric modulator of EV and EV_{back} to each. As in GLM2, we found no effect of congruency; no voxel survived when contrasting the congruency onsets nor their EV_{back} modulators. Only the contrast Congruent_{EV}<Incongruent_{EV} revealed a weak cluster in the right visual cortex (peak 38,-80,16, p<0.005 not presented). In GLM6 we split the onsets of the 1D and 2D trials by levels of EV and the 2D trials further by Congruency. No Congruency main effect survived correction. Only when the onsets of Congruent and Incongruent 2D trials with EV=70 were contrasted, a

cluster in the primary motor cortex was found (also at $p < .005$). Unsurprisingly, this cluster largely overlapped with the Congruency $\times$ EV$_{\text{back}}$ effect reported in the Main Text. Except the contrast of 1D > Congruent (see Main Text) none of the other contrasts shown in the table revealed any cluster, even at $p < .005$. **c.** Variance Inflation Factor (VIF) of the different regressors. None of the regressors (x axis) had a mean VIF value (y axis) across blocks and participants above the threshold of 4. Regressors involved in GLMs 1-4 shown on the left; GLM5 and GLM6 are shown in the middle and on the right, respectively. See Online Methods for details. **d.** Overlap of effects of EV$_{\text{back}}$ and trial type (2D > 1D). Main effects of EV$_{\text{back}}$<0 (GLM2, $p < 0.001$ FDR cluster corrected, left, blue shades) and EV$_{\text{back}}$ X Congruency < 0 (GLM3, $p < 0.005$, FDR cluster corrected, right, blue shades, t values) did not overlap with the 2D network (red shades in both panels, t values). **e.** Main effect of 1D > 2D. A stronger signal in vmPFC for 1D over 2D trials revealed weak activation in a PFC network ($p < .005$, red shades,t values). This included the vmPFC (our functional ROI is depicted in green). **f.** Stronger signal in vmPFC for 1D over congruent but not incongruent trials. When we split the onset of the 2D into Congruent and Incongruent trials (GLM5), we found no significant cluster for the 1D > Incongruent contrast, but an overlapping and stronger cluster for the 1D > Congruent contrast ($p < .001$, FDR cluster corrected, red shades, t values). We found very similar results when contrasting the onsets of 1D and Congruent in GLM6 (not presented), confirming the same results also when controlling for the number of trials for each level of EV (i.e. 1D$_{30}$+1D$_{50}$+1D$_{70}$> Congruent$_{30}$+Congruent$_{50}$+Congruent$_{70}$). Our functional ROI is depicted in green.

Table S1: Detailed univariate results: Clusters for whole brain univariate analysis, related to Fig. 8. Presented are the closest labels to the local maxima of each cluster and each contrast using AAL3v1 [84–86]. All contrasts are FDR cluster corrected. p and k values presented for each cluster.

Anatomical region			**Peak (MNI)**				**peak**	
	Label	Distance	X	Y	Z	Cluster size	t$_34$	p$_unc$
EV$_{1D}$ $>$ 0 $\cap$ EV$_{2D}$ >0 , p<001, k = 280								
R	Inferior Temporal Gyrus	4.90	60	-18	-14	1770	6.53	< .0001
R	Middle Temporal Gyrus	0	50	-6	-20		5.49	< .0001
R	Middle Temporal Gyrus	0	56	-30	-8		5.27	< .0001
R	Superior Frontal Gyrus, medial Orbital	0	8	68	-12	1045	6.09	< .0001
L	Inferior Frontal Gyrus pars orbitalis	0	-50	30	-10		4.67	< .0001
L	Superior Frontal Gyrus	0	-24	58	-6		4.35	< .0001
L	Middle Temporal Gyrus	0	-60	-30	-6	1318	5.85	< .0001
L	Middle Temporal Gyrus	0	-66	-24	-8		5.78	< .0001
L	Hippocampus	2	-40	-26	-12		4.96	< .0001
L	Angular Gyrus	0	-50	-60	38	875	5.58	< .0001
L	Angular Gyrus	0	-46	-52	30		4.86	< .0001
L	Angular Gyrus	0	-46	-70	34		3.66	.0002
L	Middle Cingulate & Paracingulate Gyri	0	-4	-40	44	1065	5.51	< .0001
L	Posterior Cingulate Gyrus	0	0	-44	32		4.52	< .0001
R	Middle Cingulate & Paracingulate Gyri	0	12	-48	32		4.52	< .0001
L	Hippocampus	0	-18	-6	-20	280	4.59	< .0001
L	Olfactory Cortex	2	-10	6	-18		4.34	< .0001
R	Angular Gyrus	0	50	-56	30	474	4.27	< .0001
R	Superior Temporal Gyrus	0	62	-54	22		4.26	< .0001
2D $>$ 1D, p<.001, k=158								
L	Superior Occipital Gyrus	2.83	-28	-76	38	5367	8.71	< .0001
L	Inferior Occipital Gyrus	0	-48	-76	-4		7.69	< .0001
L	Superior Parietal Gyrus	0	-28	-66	52		7.62	< .0001
L	Precentral Gyrus	0	-46	4	30	1766	7.69	< .0001
L	Inferior Frontal Gyrus, triangular part	0	-44	34	22		5.88	< .0001
L	Inferior Frontal Gyrus, triangular part	0	-40	26	22		5.59	< .0001
R	Inferior Parietal Gyrus	0	32	-56	54	3876	7.23	< .0001
R	Fusiform Gyrus	0	30	-76	-10		7.16	< .0001
R	Inferior Temporal Gyrus	0	48	-70	-8		7.13	< .0001
R	Inferior Frontal Gyrus, triangular part	0	48	26	26	616	5.17	< .0001
R	Precentral Gyrus	0	48	8	32		4.50	< .0001
R	Precentral Gyrus	0	38	2	30		4.23	.0001
L	Supplementary Motor Area	0	-8	14	50	159	4.69	< .0001
EV$_{m}athrmback$<0, p<.001, k = 240								
L	SupraMarginal Gyrus	2	-62	-38	22	240	4.50	< .0001
L	Superior Temporal Gyrus	0	-60	-32	10		4.26	.0001
L	Superior Temporal Gyrus	0	-60	-22	8		3.71	.0004
Congruency $\times$ EV$_{back}$<0, p<.005, k=632								
L	Postcentral Gyrus	6.93	-36	-18	60	632	4.03	.0002
L	Postcentral Gyrus	0	-48	-22	52		3.11	.0019
L	Postcentral Gyrus	0	-24	-20	74		3.08	.0020
EV$_{1D}$ + EV$_{2D}$ >0, within functional ROI, p<.001, k=979								
R	Anterior Orbital Gyrus	4.47	8	68	-12	979	7.89	< .0001
L	Superior Frontal Gyrus, Medial Orbital	2	-6	68	-12		6.86	< .0001
L	Superior Frontal Gyrus, Medial	0	-10	64	2		5.86	< .0001