Modality-specific and modality-general representations of reward value in frontal cortex

Shilpa Dang[1], Jesssica Emily Antono[1], Igor Kagan[2], Arezoo Pooresmaeili[1]*

[1] *Perception and Cognition Lab, European Neuroscience Institute Goettingen- A Joint Initiative of the University Medical Center Goettingen and the Max-Planck-Society, Germany, Grisebachstrasse 5, 37077 Goettingen, Germany*

[2] *Decision and Awareness Group, Cognitive Neuroscience Laboratory, German Primate Center (DPZ), Goettingen, Germany*

* Corresponding author: arezoo.pooresmaeili@gmail.com

Short title: Distinct neural codes for auditory and visual reward values

Number of pages: 38

Number of Figures: 5

Number of Tables: 2

Number of pages of the Supplementary Information: 8

Number of Supplementary Figures: 5

Number of Supplementary Tables: 6

Conflict of interests: The authors declare no competing interests.

Abstract

Standard neuroeconomics theories state that the value of different classes of stimuli, for instance the hedonic value of food versus music, is transformed to a common reference scale that is independent of their sensory properties. However, adaptive behaviour in a multimodal and dynamic environment requires that our brain also encodes information about the sensory features of reward predicting stimuli. Whether and how a common code for value could integrate information about the sensory features of rewarding stimuli remains inadequately understood. By employing stimuli from auditory and visual modalities as reward predicting cues in a value-based decision-making task, we were able to vary the reward value and sensory modality independently and dissociate neural codes of auditory and visual rewards in frontal areas using fMRI. Univariate fMRI analysis revealed modality-specific and modality-general value representations in orbitofrontal cortex (OFC) and ventromedial prefrontal cortex (vmPFC), respectively. Crucially, modality-specific representations were highly selective as they were only activated when participants believed that the corresponding sensory modality was associated with reward and were absent when the task involved instruction-based rather than value-based choices. Moreover, we show that modality-specific value representations are supported by the presence of the effective connectivity between each primary sensory area and the corresponding OFC activation and further between modality-specific value representations in OFC and vmPFC, only when the sensory modality to be chosen is associated with reward and absent otherwise. Our results indicate the presence of both modality-specific and modality-general representations of reward value and reveal mechanisms through which the interaction between the sensory cortices and the two types of representation guides value-based decisions.

Keywords: value-based decision-making; reward value; fMRI; orbitofrontal cortex; reward; ventromedial prefrontal cortex; sensory modality;

Introduction

When we are presented with options for making a choice, our current goals guide our decisions. Theoretical frameworks of value-based decision-making (VBDM) suggest that depending on the current goal requirements, our brain associates a subjective value to each available reward option (i.e., a valuation process), then compares these values, and makes a final choice (Balleine and Dickinson, 1998; Valentin et al., 2007; Rangel et al., 2008; Mannella et al., 2016; Eryilmaz et al., 2017; O'Doherty et al., 2017). In a multimodal dynamic environment, reward options can have fundamentally distinct sensory features, as for instance the sound of a coffee machine, the smell of fresh bread, and the sight of a bottle of our favorite smoothie in the fridge could all evoke the pleasant expectation of a nice breakfast and may therefore have the same value for us as we wake up in the morning. In addition, in real life situations value associations of stimuli can change frequently, for instance after satiation we may not enjoy the smell of bread as much as we may still be pleased by the smell of coffee.

In order to solve the choice problem in an environment as exemplified above, the valuation process in the brain should follow two principles. Firstly, it is important to be able to compare and choose between distinct stimulus options, hence stimulus-value representations independent of sensory-specific features of rewards should exist in the brain (Hampton et al., 2006; Chib et al., 2009; Hare et al., 2010; Levy and Glimcher, 2011; Noonan et al., 2011; Lin et al., 2012). Secondly, it is essential that the encoding of the most recent stimulus-value associations also includes sensory-specific information separately for each available option (Howard and Kahnt, 2021). This is important as in real world reward information can come through any sensory modality and our goals or states may change in time, requiring generation of specific predictive signals about imminent goals (e.g. whether to approach the coffee machine, the oven or the fridge in the example above) for guidance of the adaptive behavior (Klein-Flügge et al., 2013; Rudebeck and Murray, 2014; Stalnaker et al., 2014; Wilson et al., 2014; Nogueira et al., 2017). Past research has predominantly focused on the first process, where valuation in the brain adheres to a common currency coding scheme that encodes the abstract amount of the associated value independent of the identity and sensory properties of stimuli (Montague and Berns, 2002; Padoa-Schioppa and Assad, 2006, 2008; Levy and Glimcher, 2012; Berridge and Kringelbach, 2015). More recently, evidence for the second process, i.e. the identity-specific representation of reward value has been provided by studies across a range of techniques and species (for a review see Howard and Kahnt, 2021). For instance, it has been shown that specific representations exist for different flavours (McNamee

et al., 2013; Cai and Padoa-Schioppa, 2014; Suzuki et al., 2017; Howard and Kahnt, 2018) or odours (Howard et al., 2015) of juice or food items that have the same appetitive value for participants. However, the extent to which valuation follows a common currency or identity-specific coding principle and the functional significance of each coding scheme have remained unknown.

Orbitofrontal frontal cortex (OFC) and ventromedial prefrontal cortex (vmPFC) are key brain areas involved in the computation of subjective value and guidance of the value-based choices (Rolls, 2000; Montague and Berns, 2002; Rangel et al., 2008; Padoa-Schioppa, 2011; Wallis, 2012; O'Doherty, 2014; Stalnaker et al., 2015; Setogawa et al., 2019). Previous studies on human and non-human primates (Gallagher et al., 1999; Baxter et al., 2000) have shown that neurons in OFC are responsible for assigning (Noonan et al., 2010, 2011) and updating the value of individual stimulus options (Rudebeck and Murray, 2011; Rudebeck et al., 2017), for instance during the devaluation of previously rewarding events (Pickens et al., 2003), and that stimulus value signals encoded in OFC are independent of the actual final choice (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006, 2008; Kennerley et al., 2009). On the other hand, vmPFC is important in the final value-based choices (Noonan et al., 2010, 2011; Rudebeck and Murray, 2011) and its lesions impair the reward-driven decisions (Izquierdo et al., 2004; Noonan et al., 2011; Hiser and Koenigs, 2018). Importantly, different lines of evidence have pointed to the potential role of OFC, in particular the lateral OFC, in identity-specific valuation (Klein-Flügge et al., 2013; McNamee et al., 2013; Stalnaker et al., 2014; Howard et al., 2015; Howard and Kahnt, 2017), whereas vmPFC has been shown to underlie common currency coding of reward value in which different reward domains and categories, ranging from goods to monetary and social rewards, have overlapping representations (Hare et al., 2008, 2010, 2011; Chib et al., 2009; Lebreton et al., 2009; Rolls et al., 2010; Smith et al., 2010; Levy and Glimcher, 2011; Lin et al., 2012; McNamee et al., 2013). These findings raise the possibility that identity-specific and identity-general value representations might both co-exist in lateral OFC and vmPFC, respectively. In fact, both encoding of reward value (McNamee et al., 2013) and emotional valence (Čeko et al., 2022) have been shown to rely on co-existing identity-general and identity-specific representations. Whether or not the same principle extends to representation of value of stimuli from different sensory modalities and most importantly the mechanisms that generate each type of representation have remained underexplored.

Lateral and posterior regions of orbitofrontal cortex receive highly specific and non-overlapping sensory afferent inputs from auditory and visual sensory areas (Barbas, 1988,

1993; Carmichael and Price, 1996; for white matter connections - Burks et al., 2018; Martínez-Molina et al., 2019). More medial prefrontal areas including vmPFC on the other hand receive few direct sensory inputs and are more heavily connected with the limbic and visceromotor areas (Carmichael and Price, 1996). Such an organization makes these areas ideal hubs for coding reward value in a modality-specific or modality-general manner, respectively. Moreover, past research has shown that reward value modulates early sensory processing (Rutkowski and Weinberger, 2005; Shuler and Bear, 2006; Pleger et al., 2008; Serences, 2008; Goltstein et al., 2013). Together, these findings raise the possibility that to execute goal-directed choices, sensory areas communicate the information related to the identity of rewarded stimuli to the higher valuation areas such as the lateral OFC and in turn receive modulatory inputs related to the changes in goals and value structures of the environments. In the current study, we test whether such a putative mechanism can underlie value-based decision making in a dynamic multimodal environment. Specifically, we aimed to find whether and how modality-related stimulus value representations (SVR) exist in the key valuation regions when trial-by-trial updating of computed values of each sensory modality is necessary. We hypothesized that the representation of each option's value should exist in OFC in a modality-specific manner and in vmPFC in a modality-general manner, and that the co-existence of these coding schemes enables an efficient implementation of value-based choices through long-range interactions with the sensory cortices.

In order to test these hypotheses, we acquired fMRI data in a value-based decision-making task with a dynamic foraging paradigm adopted from a previous study (Serences, 2008), where subjects aimed towards maximizing their amount of monetary gain by choosing one of the two presented stimulus options which they believed was associated with a reward based on the trial-by-trial history of reward feedbacks. The two options were rewarded in an independent and random fashion to simulate foraging behavior in a varying environment. To test the influence of sensory modality through which reward information was delivered, the task was performed under three different conditions: auditory, visual, and audio-visual, where the choice was made either intra-modally (between options from same sensory domain) or inter-modally (between options form different sensory domains). To test the hypothesis that modality-specific representations in frontal areas were due to a difference in value processing requirements and not due to the difference in sensory processing requirements of the auditory and visual domains, a control task was also employed. The control task was designed in a way that the sensory processing requirements were exactly similar to the value task but selection was based on passively following an instruction as to which stimulus to choose and not on the assessment of

options' reward history. Univariate fMRI analyses revealed modality-specific and modality-general value representations in lateral-posterior OFC and vmPFC, respectively, for both sensory domains (i.e. auditory and visual). Effective connectivity analysis of a network consisting of regions exhibiting value modulations in auditory and visual sensory cortices, lateral and posterior OFC, and vmPFC, revealed how the interplay between the sensory cortices and the two types of value representations generates modality-specific representations of value and guides value-based decisions.

Materials and Methods

Participants

Twenty-four healthy subjects (13 male and 11 female, age 19 to 45 years; mean ± SD age = 27.92 ± 6.04 years) participated in the experiment for financial compensation of 8€/hour. The sample size was based on a previous study that used a similar paradigm (Serences, 2008, N = 15) but to account for possible dropouts we tested N=24. The experiment was done in two scanning sessions, each lasting about 2.75 hours, preceded by one online training session (0.5 hours) to familiarise the participants with the task. Participants also had the opportunity to earn a monetary bonus of maximum 22€ based on their behavioural performance in the value-based decision-making task (value task) during the scanning session. All participants were right-handed and had normal or corrected-to-normal vision, and were naïve to the hypothesis of the project. Before the experiment started and after all procedures were explained, participants gave an informed written consent and participated in a practice session. The study was approved by the local ethics committee of the "Universitätsmedizin Göttingen" (UMG), under the proposal number 15/7/15.

Four participants were excluded from the final analysis resulting in the data from 20 subjects presented here: two participants had difficulty in differentiating the strategies of the value and the control task (specifically with the instructions associated with the feedback colours in the two tasks, see the Experimental Design); one participant was excluded due to excessive head motion while scanning (> 4 mm); and one participant due to the unusually large size of the ventricles in the structural MRI scan (judged by a co-author who had training in medicine).

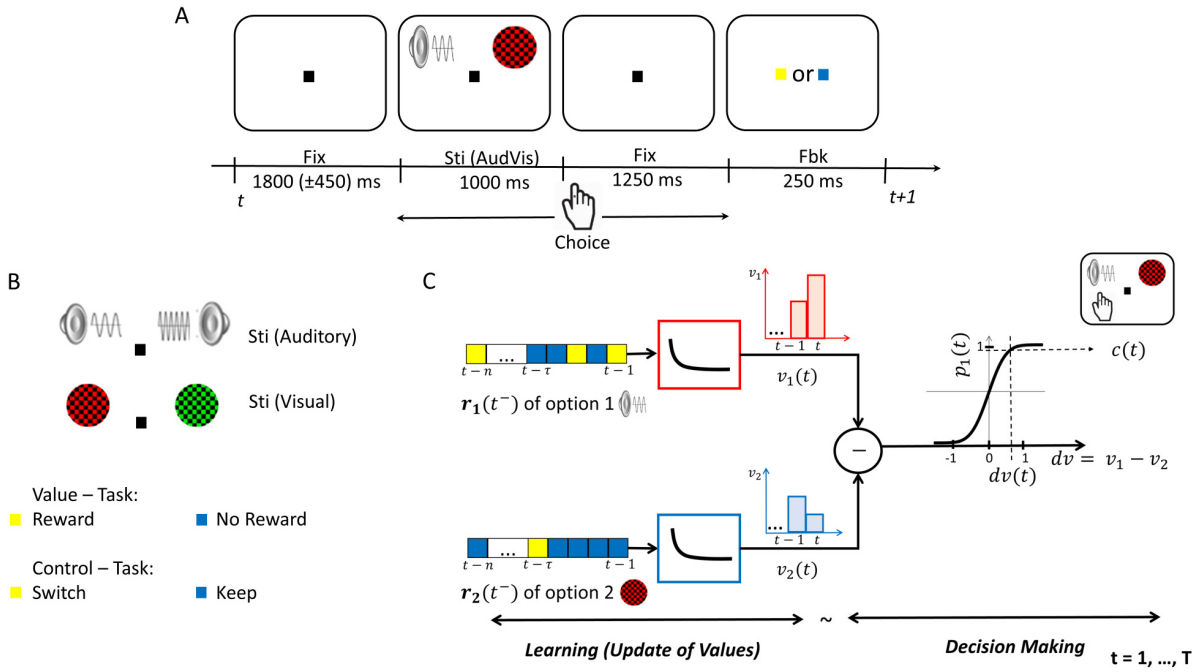

Figure 1. Experimental Paradigm and Computational Framework of Choice Behaviour. (**A**) General schematic of an audio-visual (AudVis) trial across both behavioural tasks (i.e., value and control tasks). After a jittered inter-trial interval of fixation (Fix), stimuli (Sti) options were presented. Participants made a choice during a response window of fixed interval (i.e., 2.25 s from the onset of stimuli), after which the fixation changed to either yellow or blue colour to indicate the feedback (Fbk). (**B**) Stimuli options as presented during an auditory or a visual trial. In the value-based decision-making task, the yellow feedback indicated a reward and blue feedback no reward, whereas in the control task, the yellow and blue feedback instructed participants either to switch or to keep the past trial's choice, respectively. (**C**) Reward history of option 1 and 2; i.e. $r_1(t^-)$ and $r_2(t^-)$; enter as inputs to two identical exponentially decaying filters that weigh rewards based on their time in the past and compute the subjective value of each option (i.e. v_1 and v_2). The difference of the output of filters gives the differential value between the options (i.e., dv). The differential value according to a sigmoidal decision criterion results into the probability of choice (option 1 or 2, here option 1 is chosen as an example, see *equation 5* in the text).

Experimental design

The experiment consisted of a value-based (value task) and an instruction-based (control task) decision-making task, completed in two sessions (**Figure 1**). Each session consisted of 12 blocks (of 72 trials each): 9 blocks of the value task followed by 3 blocks of the control task. Each of the tasks involved a binary choice between stimuli presented in three sensory domains: both auditory (*AudAud*), both visual (*VisVis*), and audio-visual (*AudVis*), which were presented in separate blocks. All three types of sensory domain blocks appeared an equal number of times across each task in a pseudo-random order.

Stimuli: Two pure auditory tones (low pitch (LP) tone- sawtooth, 294 Hz; high pitch (HP) tone- sinusoidal, 1000 Hz, played through MR-compatible earphones -Sensimetric S15, Sensimetrics Corporation, Gloucester, MA- with an eartip -Comply™ Foam Canal Tips-) and two contrast reversing visual checkerboards (green and black; red and black, as in (Serences, 2008)) within circular apertures (4˚ radius) were used as the choice options. In an auditory

(*AudAud*) or a visual (*VisVis*) trial, either two tones or two checkerboards were presented as options, respectively. In an audio-visual (*AudVis*) trial, one tone and one checkerboard were presented as options. Choice options were presented simultaneously on the left or right side of the centre (for auditory tones to each ear). All combinations of a tone and a coloured checkerboard: LP-Red, LP-Green, HP-Red, HP-Green, were presented an equal number of times across the 72 trials of *AudAud*, *VisVis* and *AudVis* blocks in a pseudorandom order.

Trial Structure: Both value and control tasks were the same in terms of presentation of stimuli options, response requirements, and feedback on the decision. The only difference was over the indications associated with the feedback colours (**Figure 1A-B**). Participants were asked to fixate continuously throughout each run (here, a run = 3 blocks) on a small square (0.4˚visual degree) at the centre of the screen. A trial began with a mean fixation period of 1.8 s, yielding a mean trial duration of 4.3 s. Following the fixation period, the two stimuli options were presented simultaneously for 1 s, one on each side of fixation (left and right; auditory stimuli were played one on each side of the earphones; visual stimuli were centred 10˚to each side and 5˚above the centre of the screen). The spatial position of each option was also pseudo-randomised across the trials of a block in such a way that each option appeared an equal number of times on both sides of the fixation point. Following the onset of the stimuli options, participants pressed either the left or the right button on a MR-compatible two-buttoned response box (Current designs Inc., Philadelphia, PA), using the index or the middle finger of their right hand, to indicate their choice. The participants were required to respond within 2.25 s following the onset of the options. Following the response window, a feedback window of 0.25 s appeared in which the central fixation point turned either yellow or blue in colour. In the value task, the yellow fixation indicated that the choice was rewarded and the blue fixation indicated that the choice was not rewarded. Since the control task was designed to be similar to the value task in terms of sensory processing requirements without a need to track and update their estimation of options' value, the feedback instructed the participant to make a prespecified choice. Thus, in the control task, the yellow fixation indicated to switch from the past trial choice and the blue fixation indicated to keep the past trial choice. The choice on the first trial of any control block was a random choice.

Dynamic reward structure: To create a dynamic multimodal environment for participants, the rewards were assigned to the options from different sensory modalities independently and stochastically at random intervals using a Poisson process (Corrado et al., 2005). On average, a reward was available for delivery on 33% of the trials (of a block of value task). These 24 rewards in a block (33% of 72 trials) were distributed between the two stimuli

options in different reward ratios of {1:3, 1:1, 3:1}, such that the rewards assigned to options were {8.5%:24.5%, 16.5%:16.5%, 24.5%:8.5%} in percentage of trials. For the value task in a single session (9 blocks), these three reward ratios were repeated and randomized such that each reward ratio was used exactly once with every sensory domain block (i.e. *AudAud*, *VisVis*, and *AudVis*). The randomization of various factors such as sensory modality, spatial position of options, and reward ratios was done to provide a dynamic environment, in which the participant would be required not only to update their stimuli-value associations with changing reward ratios but also to keep track of the rewarding stimuli very carefully on a trial-to-trial basis with changing spatial positions. Two important schemes of "baiting" and "change over delay" (COD) were adopted as in previous studies (Corrado et al., 2005; Serences, 2008). In baiting, an assigned reward to an option remained available until that option was chosen. This was done to avoid the "extreme exploitation" strategy in which a participant would always stick to the option with a higher reward rate (e.g., 24.5%>8.5%) association in a block and to motivate the exploration in which a participant should visit both reward options occasionally. Also, an earned reward feedback was delayed for one trial when the participant changed their choice from one option to the other and delivered only if the participant chose the same option again. This cost, i.e., COD, was employed to discourage "extreme exploration" strategy, where the participant would be able to consume all rewards without any learning by alternating choices rapidly between options. Trials following a change of choice (switch) between options were not included in the behavioural analysis (and were marked by a specific regressor in fMRI analysis) because subjects were informed that they will not get a reward on such trials and hence choices were not completely free. At the end of each block, participants were shown the reward earned in that block at the rate of 5 cents per yellow square shown as the reward feedback. At the end of the second session, participants received the total reward earned which was up to a maximum of 22€ (11€ per session) based on their performance along with a participation fee of 8€ per hour.

Control task structure: Similar to the reward structure in the value task, switches were assigned independently and stochastically to the options in an equiprobable manner with an average switch rate of 33%. Thus, on any trial when a participant earned a switch from a chosen option, yellow feedback was displayed indicating that they should switch their choice to the other option on the next trial. On other trials, when a switch was not assigned, blue feedback was shown to indicate that the same option should be chosen on the next trial. This type of switch assignment structure was developed to encourage a similar temporal choice pattern as in the value task. On a single day, the control task was conducted in each of the three sensory

283 domains. There were no baiting and COD schemes employed in the control task. At the end of
284 each block, participants were shown their performance that indicated how accurately they
285 followed the instructions in that block.

286 *Computational framework of choice behaviour*

287 To examine whether participants' choices in the value task were influenced by the
288 dynamic reward structure employed in our design, we used a computational framework that
289 has been used in the past to model choice behaviour abiding by the matching law (Herrnstein
290 and Baum, 1970; Corrado et al., 2005; Serences, 2008). In our task, there were no prior reward
291 associations with the options, and hence on any trial t a participant made a choice $c(t)$ based
292 on the previous rewards received $r(t^-)$ during the experiment, see **Figure 1C**. Intuitively, an
293 option that delivers more rewards per unit of time should have relatively higher value
294 associations and should be chosen more often. The value associations, in general, should
295 strongly predict how a participant's choices are affected by reward history, i.e., "the learning
296 mechanism". Thus, to estimate participants' subjective value beliefs for each reward option on
297 a trial-by-trial basis, we fitted the reward history and choice data of each participant to a linear-
298 nonlinear-probabilistic (LNP) model, shown in **Figure 1C** (also called as linear regression-
299 based model of reinforcement learning (Katahira, 2015)). Two broad phases of the LNP model
300 are the *learning* and the *decision-making* (for implementation details refer to Corrado et al.,
301 2005; Serences, 2008).

302 In the *learning phase* (see **Figure 1C**), two identical linear filters (n learning
303 weights $\alpha_\tau, \tau = 1\ to\ n$ trials in the past) weigh the reward history of each option ($r_i(t^-), i =$
304 1,2 correspond to stimulus sets S_1, S_2) based on the reward received on each of them in past.
305 The value for n was taken to be half of the trials over which the reward ratio was unchanged
306 (here, n = 36), as has been done before (Serences, 2008). This results in updating of the value
307 belief of the options, $v_i(t), i = 1,2$. As the overall reward assignment over the two options was
308 symmetric, their impact on choice was equal and opposite, hence the linear filter was derived
309 by closely matching the composite reward history $\boldsymbol{r}$ (as shown in (1)) and composite choices $\boldsymbol{c}$
310 (shown in (2)).

311 $$\boldsymbol{r} = \boldsymbol{r_1} - \boldsymbol{r_2} \qquad \textbf{(1)}$$

312 $$\boldsymbol{c} = \boldsymbol{c_1} - \boldsymbol{c_2} \qquad \textbf{(2)}$$

313

314 The *decision-making phase* (see **Figure 1C**), draws the ultimate binary choice (S_1 or
315 S_2) on trial t based on a relation that maps the differential value $dv(t)$ (as shown in (3))

computed on trial t to the participant's probability of choosing option S_1 on that trial. Intuitively, this relation should strongly predict a participant's choice behaviour, where the participant should make a choice $c(t)$ based on the comparison process shown in (4).

$$dv(t) = v_1(t) - v_2(t) \quad \textbf{(3)}$$

$$c(t) = \begin{cases} S_1, if\, v_1 > v_2 \\ S_2, if\, v_1 < v_2 \end{cases} \quad \textbf{(4)}$$

To assess the fit of the LNP model during the *learning phase*, the linear filter weights for the data of each participant were approximated by fitting an exponentially decaying function, indicating that choices were most impacted by recent rewards rather than distant rewards in past (quantified by time scale parameter τ trials of the fit; see **Figure 1C** for the illustration of the filter and **Figure 2C** for the fit to the data of a single participant). To assess the *decision-making phase*, the probability of choosing option S_1 for each participant was approximated by fitting a normal cumulative distribution function (equation (5)).

$$\varphi(x, \mu, \sigma) = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad \textbf{(5)}$$

where x is the differential value (dv). This function contains two important decision-making parameters: μ corresponding to participant's biasness towards a particular option and σ that measures the sensitivity to value differences or in other words the explore-exploit tendency. Accordingly, $\sigma = 0$ corresponds to an extreme exploitative tendency, and $\sigma \to \infty$ to extreme exploration. The disadvantage of being extremely exploitative; i.e., sticking to an option that has higher reward rate associated with it, is that it would yield lesser number of rewards to the participant because there exist unvisited options, which remain baited until chosen. Moreover, extreme exploration would also be disadvantageous, as it would lead to no learning and the absence of any strategy. Thus, the optimal strategy in this task would be to choose more often the option with higher reward rate and to occasionally visit the less rewarding option to consume rewards on it. An optimal strategy is advantageous in a dynamic reward structure task where the aim is to maximize rewards, and to examine whether this is the case in our task, we inspected the abovementioned parameters (τ, μ, σ) for their fit to participants' behavioural data (**Figure 2C-E**).

In the value task, the positive and negative feedbacks have distinct effects on participants' beliefs. Therefore, if the choice of a particular option was rewarded (yellow feedback) or not rewarded (blue feedback) on the previous trial, then the value beliefs for that option should be relatively higher or lower, respectively, on the current trial in comparison to the value beliefs in the past trial. As only one of the two options could be chosen and rewarded

in any trial, the differential value of two options would also be relatively high in magnitude when reward was received on the past trial, and otherwise low. On the contrary, the control task was designed in a way to be like the value task in terms of sensory processing requirements but not involve the participant in any learning or updating of the stimuli value. Intuitively, when no learning via feedbacks occurs, the two types of feedbacks (keep/switch) should have a similar effect on the subjective preference over the options. To confirm this, we tested the fit of the same LNP model to the choices in the control task and compared the absolute differential values of each trial obtained from models' fits to both tasks (value and control tasks) against the type of feedback received (blue or yellow) in the previous trial (**Figure 2F**). We used the absolute differential values (*absDVs*) as a measure of subjective preferences because the choice behaviour is symmetric with respect to the individual options.

fMRI data acquisition and pre-processing

MRI scanning was carried out on a 3-Tesla Siemens MAGNETOM Prisma scanner equipped with a 64-channel head-neck coil at the Universitätsmedizin Göttingen. Anatomical images were acquired using an MPRAGE T1-weighted sequence that yielded images with a 1 x 1 x 1 mm resolution. Whole-brain multi-shot Echoplanar imaging (EPI) volumes were acquired in 69 interleaved transverse slices (TR = 1500 ms, TE = 30 ms, flip angle = 70°, image matrix = 104 x 104, field of view = 210 mm, slice thickness = 2 mm, 0.2 mm gap, PE acceleration factor = 3, GRAPPA factor = 2). Data from each participant was collected in two identical sessions on two separate days. An experimental session consisted of multiple runs of fMRI data acquisition, where a run comprised starting the scan and acquiring data for three blocks of the tasks (~ 20 minutes) after which the scan was stopped and resumed again after a break (~ five minutes). On each day, four fMRI runs (first three runs: 9 blocks of the value task, last run: 3 blocks of the control task) were conducted and each fMRI run lasted 16.355 min.

Data pre-processing and further statistical analyses were performed using Statistical Parametric Mapping software (version SPM12: v7487; https://www.fil.ion.ucl.ac.uk/spm/) and custom time-series analysis routines written in MATLAB. EPI images of each session were slice time corrected, motion corrected, and distortion corrected by using the measured field maps. The T1 anatomical image was co-registered to the mean EPI from realign-&-unwarp step, and then segmented. The estimated deformation fields from the segmentation were used for spatial normalization of the corrected functional and anatomical images from the native to

the MNI (Montreal Neurological Institute) space. Finally, the normalised EPI images were spatially smoothed using a 6 x 6 x 6 mm FWHM Gaussian kernel.

fMRI univariate analysis: General linear modelling (univarGLM)

For each participant, we first specified a general linear model (GLM) using the pre-processed functional images of two sessions that were concatenated one after another. The GLM modelled both the value and the control task using 35 event-related regressors convolved with the canonical hemodynamic response function (HRF). For the value task, we defined individually for each of the three modality conditions (auditory, visual, audio-visual) one unmodulated stick regressor representing the modality-wise trial identity and two parametrically modulated stick regressors containing the trial-by-trial updated subjective value (SV) beliefs regarding each of the options presented, referred to as the value-modulated regressors. Trial identity was entered as 1 at the onset of the stimuli for trials of a particular modality condition and 0 otherwise. The value-modulated regressors represent the trial-by-trial learning and updating of value beliefs for each option separately (**Figure1C and Supplementary Figure S2**), and are denoted by *lpSV* and *hpSV* in auditory domain corresponding to low pitch and high pitch tones, *rSV* and *gSV* in visual domain corresponding to red and green checkerboard stimuli, *aSV* and *vSV* in audio-visual domain corresponding to auditory and visual stimuli in any combination (see also **Table 1**). The trial-by-trial SVs were entered at the onset of the stimuli options.

Similarly, for the control task we defined individually for each of the three modality-domains one unmodulated regressor representing the modality-wise trial identity and two parametrically modulated regressors corresponding to each of the options presented. In the control task, the aim was to passively follow instructions. Thus, to create a parametrically modulated regressor corresponding to one stimulus option, a weight of either 1 or 0 was assigned at the onset of stimuli options in each trial depending on whether the instruction (keep/switch your choice) from the last trial was correctly followed or not, respectively (see also **Supplementary Table S1** and **Figure S2**).

We also included two unmodulated event-related regressors (collapsed across the value and the control task) locked to the time of response and the onset of feedback. 15 nuisance regressors were included corresponding to the following: instruction presentation at the start of each block, six motion parameters, run regressors (modelled by assigning a weight of 1 for each volume of that run and else 0: a run corresponds to each period of MRI data acquisition between the start and the end of the scan) to account for the difference in the mean signal

activity between each time the scan started (one less in number than the total number of fMRI runs, here 7) and a constant.

To identify the neural correlates of modality-related stimulus value representations (SVR), we contrasted parameter estimates of value-modulated regressors against baseline separately for each sensory domain (for definition of contrasts see **Table 1**). Note that our primary interest in this study was to identify the neural correlates of valuation for different sensory configurations. Since in a binary choice situation valuation occurs for each of the two options separately, we used the sum of estimated responses to each option as our dependent variable. This is a different approach than using the differential value of the options as the dependent variable which capitalizes on identifying the neural correlates of comparison and choice between options (Serences, 2008) rather than the valuation of each individual option. On account of previous studies identifying the domain-general and domain-specific valuation areas in the vmPFC and OFC (McNamee et al., 2013; Howard et al., 2015; Howard and Kahnt, 2017), we limited our analysis to a mask encompassing the orbital surface of frontal gyrus. This search volume (for details see **Figure S1**) consisted of anatomical parcellations of orbital surface of frontal gyrus as defined in automated anatomical labelling (AAL) atlas (Rolls et al., 2015, 2020). Statistical maps were assessed for cluster-wise significance using a cluster-defining threshold of $t(19) = 3.58$, $P = 0.001$; and using small volume corrected threshold of $P < 0.005$ (referred to as a small volume family-wise-error (SVFWE) correction) within the frontal search volume. Whole-brain results were inspected at FWE $p<0.05$, and $k>10$ (see **Table 2**).

Effective connectivity analysis of fMRI data

Our univariate analysis identified a number of regions, both in sensory and in frontal areas, that were modulated by individual stimulus values computed prior to a choice was made (**Figure 3, Table 2** and **Supplementary Figure S5**). We next aimed to determine whether and how the long-range communication between these areas generates the modality-specific representations of value and guides the final choice. To this end, we investigated the modality-specific effective connectivity (EC) of a network consisting of sensory and frontal regions exhibiting value modulations at the time of options' presentation by employing deterministic bilinear dynamic causal modelling (DCM) approach (Friston et al., 2003; Stephan et al., 2009; Friston, 2011). This approach fits a set of pre-defined patterns of EC within a model space to the fMRI time series and compares them in terms of their evidence (for details of the model space see the section under *Defining the model space for a 5-node network*).

Modality-specific effective connectivity is the connectivity among sensory and frontal regions during the valuation of a particular modality condition, either auditory or visual. For intramodal conditions (*AudAud*, *VisVis*), the EC of the network is clearly linked to one single modality, either auditory or visual (denoted by *intraAud* or *intraVis,* respectively). However, for intermodal condition (*AudVis*), changes in EC during the valuation process occur for both auditory and visual modalities. To achieve the maximum separation between modalities during the intermodal condition which would then allow us not only to determine modality-specific EC in full network but also to test the same models for their fit to both intra- and inter-modal conditions, we separated trials in the intermodal condition according to whether the auditory or visual stimulus was selected (hence denoted by *interAud* or *interVis*, respectively). Using the same model space for both intra- and inter-modal conditions will provide insights on the underlying mechanisms that mediate the modality-specific valuation across different contexts, i.e., when the same or different sensory modalities are compared against each other in terms of their value. Additionally, this approach is more parsimonious than either having two separate sets of models for each condition or increasing the number or complexity of the models to account for the differences between inter- and inter-modal conditions (Vandekerckhove et al., 2015). Accordingly, to determine modality-specific effective connectivity, we estimated an additional GLM. This GLM was identical to that used for univariate analysis (univarGLM), except that for both tasks (the value and the control task), intermodal condition was separated into auditory and visual trials on the basis of the final choice.

Regions of interest (ROIs): ROIs for the effective connectivity analysis comprised the frontal valuation areas and the sensory regions that contained stimulus value representations for auditory and visual modalities according to the univariate analysis (**Figure 3, Table 2** and **Supplementary Figure S5**). The resulting five ROIs from which the representative time series for DCM analysis were extracted were as follows: 1) The overlapping activation area for visual and auditory value representations in vmPFC during each respective condition, 2) The overlapping activation area of left latpostOFC during intra-modal auditory and inter-modal auditory conditions – i.e., audOFC, 3) The overlapping activation area of left latpostOFC during intra-modal visual and inter-modal visual conditions – i.e., visOFC, 4) bilateral activations in auditory sensory cortex – i.e., audSen, and 5) bilateral activations in visual sensory cortex – i.e, visSen. The representative time series for any ROI was the first principal component of the pre-processed fMRI time series of the selected ROI.

Defining the model space for a 5-node network: In order to understand how modality-specific valuation is supported by a network comprising modality-general and

modality-specific areas, we estimated 21 biologically plausible models for the value and the control tasks, with three types of connections: driving input, intrinsic, and modulatory. These models were developed over a base model comprising driving inputs and intrinsic connections, which did not vary with the experimental conditions. The models differed from each other over modulatory connections, which depended on the experimental conditions. In the base model, intrinsic connections were defined between every pair of nodes in the network and as self-connections. Because the stimuli were presented aurally or visually, two types of driving inputs to the network were defined for auditory and visual sensory cortices: 1) an input to ROI audSen in auditory and audio-visual conditions of both tasks, and 2) an input to ROI visSen in visual and audio-visual conditions of both tasks. The driving input was modelled by entering ones at the onset of stimuli options belonging to a certain condition type and else zeros (see also the **Supplementary Tables S5 and S6**).

The model space of all possible connectivity models would be extensive for a 5-node network (Friston et al., 2011), where a modulatory connection between any two nodes can exist in none or more of the 4 experimental conditions of a task (*intraAud*, *intraVis*, *interAud*, *interVis*) and 2 directions (directed and reciprocal). Thus, we constrained the model space based on the following assumptions:

1) We included models with only bidirectional modulatory connections between nodes (Friston et al., 2011), based on the past findings that anatomical connectivity between two cortical areas is generally bidirectional (Zeki and Shipp, 1988). Additionally, large connectivity databases indicate a strong likelihood of cortico-cortical connections to be reciprocal (Kötter and Stephan, 2003). Moreover, this constraint does not imply that connection strengths would be identical for both the directed and reciprocal connection between two nodes. Further, we included models with each node connected (modulatory) to at least one other node of the network with the exception of a null model, which has no modulatory connectivity in the network due to experimental conditions of any task.

2) We observed from univariate analysis results in the intra-modal condition that the value activations in ROIs audOFC and visOFC were mutually exclusive, thus we included models with no modulatory connection between these two nodes. We note that this constraint may not apply to the data of inter-modal condition. However, to be able to use the simplest and least exhaustive model space that could be tested for its fit to the data of both intra- and inter-modal conditions, we assumed this constraint to also hold for inter-modal trials. This assumption is plausible since inter-modal trials were separated

based on the modality that determined the final choice, and the evidence accumulation process that drives the final value-based choices is most strongly influenced by the information related to the valuation of the chosen compared to the unchosen modality, especially in vmPFC (Wunderlich et al., 2012).

3) Further, we observed that auditory valuation ROIs – i.e., vmPFC, audOFC, and audSen- were significantly activated as auditory SVR, thus forming an auditory value sub-network. Similarly, the visual valuation ROIs -i.e., vmPFC, visOFC, and visSen- were significantly activated as visual SVR, forming a visual value sub-network. Thus, we included models with symmetric modulatory connections across the two sub-networks.

This resulted in a biologically plausible connectivity model space consisting of 10 models per task (shown in **Figure 4A**) plus a null model. We estimated each of the 21 models individually for all the 20 subjects. However, for one subject the parameter estimation did not converge and therefore, we excluded that subject from the effective connectivity analysis. Thereafter, we identified the most likely model using a group-level random effects Bayesian model selection (rfxBMS) approach (Stephan et al., 2009). The model exceedance probability used to find the best model as shown in **Figure 4B** represents the probability that a particular model m is more likely than any other model in the model space (comprising of M models), given the group data. Note that the exceedance probabilities over the model space add to one (Stephan et al., 2009). Next, we estimated the connection strength parameters for connections of interest using Bayesian parameter averaging (BPA) approach (**Figure 4C**).

Results

Behavioural results

We examined participants' performance in two behavioural tasks (**Figure 1**) referred to as the value-based (value) and the instruction-based (control) tasks. In both tasks, participants aimed at maximizing their performance, i.e., the reward magnitude in the value task and the accuracy in following the instructions in the control task, by selecting one of the two presented stimuli options (either two auditory stimuli or two visual stimuli or one auditory and one visual stimulus, as shown in **Figure 1A-B**; for details see Material and Methods). A choice was made either from stimulus set S_1 = {low pitch, green, auditory} or corresponding stimulus set S_2 = {high pitch, red, visual}.

In the value task, participants experienced an unpredictable outcome scenario with a dynamic reward structure (see Materials and Methods). Reward baiting and change over delay (COD) strategies along with uninformed changes in the reward ratio across every block of trials motivated an exploratory choice pattern. Overall, the choice pattern in the value task exhibited matching behaviour nearly in accordance with the Herrnstein's Matching Law (Herrnstein and Baum, 1970), which relates the choice behaviour to reward ratios {1:3, 1:1, 3:1}, as shown in **Figure 2A** for all modality domains (auditory, visual, audio-visual). Specifically, the choice ratios, which indicate the number of choices made towards one reward option (S_1) over another (S_2), increased as the $S_1{:}S_2$ reward ratio increased. Importantly, the choice patterns were consistent across sensory modalities. This effect was captured by a strong main effect of reward ratio on choice ratio ($F[2,38] = 183.8$, $p < 0.001$) and no significant interaction between reward ratios and options' modality ($F[4,171] = 0.95$, $p = 0.34$) in a two-way repeated-measures ANOVA. Only a weak effect of modality on choice ratios was observed ($F[2,38] = 5.95$, $p = 0.024$), which corresponded to a tendency of participants to choose visual option more often than auditory options in the audio-visual block (for more details see the Supplementary Information). Therefore, we collapsed the behavioural analysis results across modalities for a concise presentation of results (choice ratios collapsed across modalities for each reward ratio, as shown in **Figure 2B**).

We next tested whether participants' choices in the value task followed the predictions of our computational framework; i.e., they adhered to an LNP model (for details see Material and Methods). To this end, we approximated the linear filter weights using the best-fitting exponentially decaying function (quantified by time scale parameter τ; **Figure 2C**), and the probability of choice function using the best-fitting sigmoidal function (quantified by biasness μ, and sensitivity to value differences σ; **Figure 2D**), for each participant. Across participants the mean time scale parameter τ was 1.22 (±0.15 s.e.m), which was significantly greater than zero $t[19] = 8.39$, $p < 0.05$, indicating that choices were in fact most impacted by recent rewards rather than distant rewards in the past (**Figure 2E**). Mean biasness μ across participants was 0.07 (±0.07 s.e.m), which was not significantly different than zero $t[19] = 0.66$, $p = 0.52$, indicating that participants did not have a bias towards any particular option. Finally, the mean sensitivity σ across participants was 0.81 (±0.10 s.e.m), which was significantly greater than zero $t[19] = 8.81$, $p < 0.05$ and insignificantly lesser than one $t[19] = 1.95$, $p = 0.07$, indicating that participants were aware of the value difference between options and had indeed adopted an optimal balance between exploration and exploitation. Following this optimal strategy,

participants were able to harvest 94.94% (±0.84% s.e.m) of the total rewards available. Overall, participants exhibited choice behaviours that were strongly predicted by the filter weights, estimated subjective values, and sigmoidal decision-criteria of the LNP model.

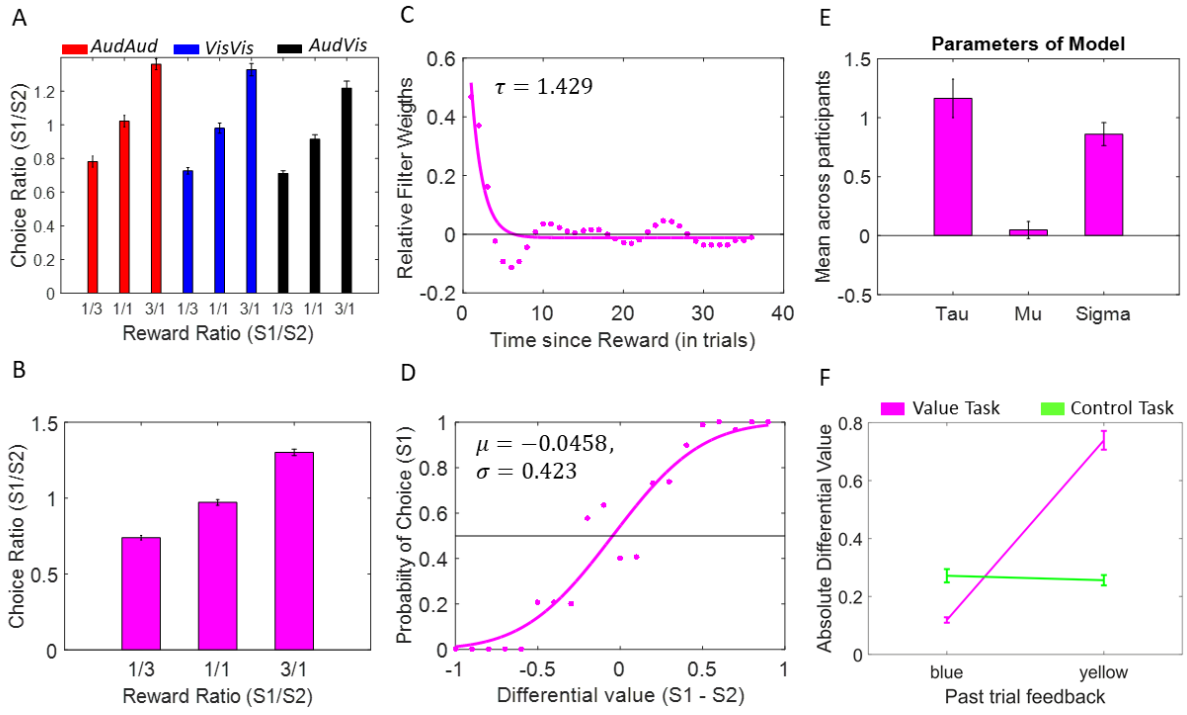

Figure 2. Behavioural Data. (A) Mean choice ratio across participants for each reward ratio {1:3, 1:1, 3:1} over option 1: option 2, separately for each modality condition of the value task (*AudAud*, *VisVis* and *AudVis*). (B) Mean $S_1{:}S_2$ choice ratios in (A) collapsed across modalities for individual reward ratios. (C) Linear filter weights (dots) and exponential approximation (solid line) showing how past rewards are weighed based on their time in past for a single participant in the value task. Parameter τ shows the timescale component of the best-fitting exponential. (D) Mapping of differential value of option 1 and 2 to the probability of choice for option 1 (dots) and sigmoidal approximation (solid line) for the same participant as in (C). Parameters μ and σ of the best-fitting cumulative normal function show the participant's biasness towards an option and sensitivity to value differences, respectively. (E) Mean parameters of the best-fitting curves across participants. (F) Relationship between feedback colors and absolute differential value for value and control task, across all participants. $S_1{:}S_2 = \{low\ pitch{:}\ high\ pitch,\ green{:}\ red, auditory{:}\ visual\ \}, S_1 - S_2 = \{low\ pitch - high\ pitch,\ green - red, auditory - visual\}$. Error bars indicate standard error of the mean (s.e.m.) across participants.

In order to demonstrate that the LNP model uniquely predicted the learning and choices in the value task, the fit parameters were also inspected for the data of the control task. Overall, in the control task participants passively followed the instruction provided by the feedbacks with a high accuracy (i.e., 95.2% ± 1.33% collapsed across keep/switch feedbacks), which indicated that they were aware of the task strategy. As choices in the control task were instructed, participants' choices in this task were expected not to reflect any trial history-based tracking of option values beyond the instruction provided in the immediately preceding trial. In fact, this is exactly what we found when we compared participants' beliefs about options' value in the current trial depending on the type of feedback received in the past trial (blue or yellow). This effect (**Figure 2F**) was captured by a significant interaction $F[1,19] = 254.7$, p <

0.001 between the task type (value or control) and feedback (yellow or blue) on determining the absolute differential values (*absDVs*; a measure of subjective preferences), where *absDVs* showed a significant difference between the two types of feedbacks in the value task (mean±s.e.m. = 0.12±0.01 and 0.74±0.03 for blue and yellow feedbacks, respectively, $p< 10^{-11}$) but not in the control task (mean±s.e.m. = 0.27±0.02 and 0.26±0.02 for blue and yellow feedbacks, respectively, p = 0. 11). Analysis of the mean reaction times (RT) for the two types of feedbacks in the value and the control tasks revealed no significant mean or interaction effect (all ps>0.05, for details see the **Supplementary Information**).

Overall, our behavioural results confirmed that in the value task participants learned and updated their beliefs about options' values through monitoring the feedbacks received on each trial, whereas in the control task they passively followed the instructions without any further processing of stimulus value, as intended.

fMRI results

Modality-general and modality-specific stimulus value representations: vmPFC and OFC

In order to identify the modality-specific and modality-general stimulus value representations in the frontal cortex (see the Material and Methods and **Figure S1** for specifications of the search area), we performed group-level random-effects analysis on the contrast images obtained from fMRI data of all participants. In intra-modal conditions (*AudAud* and *VisVis*), we estimated an overall effect of value-modulated regressors separately in auditory and visual sensory domains by defining contrasts: *intraaudSV* > 0 and *intravisSV* > 0, where *intraaudSV>0* = *lpSV>0* + *hpSV>0* and *intravisSV>0* = *rSV>0* + *gSV>0* (see **Table 1** for the detailed description of contrasts).

The auditory contrast revealed significant activations in vmPFC and left lateral OFC (latOFC) and the visual contrast activations in vmPFC and left posterior OFC (postOFC, **Figure 3**). However, we did not find any significant activation in the right OFC for either of these contrasts. Lateralization of reward responsiveness in OFC could be related to a functional specialization of the left and right lateral OFC and has been reported in the past (Lopez-Persem et al., 2020). Crucially, we found a segregation of value-processing clusters across the sensory domains in OFC (d = 20.59 mm, d: Euclidean distance). On the contrary, the auditory and visual clusters in vmPFC were substantially overlapping with a separation d < 8 mm between the cluster peaks (for separation criteria - see Poline et al., 1997; Hallett, 1998). These findings

provide answers for our study's first question of whether modality-general and modality-specific stimulus value representations exist in the valuation regions of the frontal cortex.

To determine whether and how the modality-wise segregation of value modulations in OFC persists during the simultaneous presentation of options from both sensory domains in the inter-modal condition (*AudVis*), we estimated the average effect of value modulations in both sensory modalities using the contrast: *interaudvisSV* > 0, where *interaudvisSV>0* = *aSV>0* + *vSV>0* (see **Table 1** and **Figure S3**). This contrast showed significantly activated clusters in vmPFC and in left anterior OFC (antOFC). Interestingly, we observed that in the activated latOFC cluster, there were two local maxima peaks which were closer to the auditory and visual peaks found in the intra-modal conditions (-48, 30, -12, $t(19) = 3.93$ and -36, 24, -18, $t(19) = 4.61$). Further, to specifically test whether these peaks corresponded to individual sensory modality value representations, we estimated the individual effects of value in sensory modalities in inter-modal condition by using contrasts: *aSV* > 0 and *vSV* > 0. The individual contrasts revealed overlapping clusters in vmPFC ($d = 6.63$ mm) and separate clusters in lateral and posterior OFC ($d = 18.76$ mm) (**Figure 3, Table 1**). These modality-specific valuation clusters in OFC were found to overlap with their respective modality clusters found in intra-modal conditions (separation $d < 8$ mm), as shown in **Figure 3 F-G**.

An alternative explanation for the segregation in modality-wise representations is that rather than reflecting the functional specialization of OFC neurons for the visual and auditory values, they reflect differences in the sensory properties of stimuli options. To rule out this possibility, we next examined the control task (for details see the **Supplementary Information**). Crucially, the modality-specific activations in OFC were absent in the control task when the same contrasts as in the value task were examined, demonstrating that they exclusively reflect the trial-by-trial updating of stimulus-value associations rather than the sensory features of stimuli or choice based on the instruction. On the contrary, in the control task weak activations overlapping with the modality-general representations in vmPFC were found (see **Figure S4** and **Table S3**) highlighting a general role of this area in representing the final choice irrespective of whether or not choices are informed by value or are instructed. These results were corroborated by testing more stringent interaction contrasts that compared the value and control task against each other (see the **Supplementary Information** and **Table S4**).

Table 1. Modality-related stimulus value representations in vmPFC and OFC for various contrasts

Contrast	Region	*X*	*Y*	*Z*	*t*(19)	*k*
intraaudSV > 0	vmPFC	-6	58	-14	5.33	121
	latOFC_L	-48	32	-10	5.31	134
intravisSV > 0	vmPFC	-8	58	-14	4.70	142
	postOFC_L	-30	26	-18	6.20	143
interaudvisSV > 0	vmPFC	0	52	-10	6.93	537
	antOFC_L	-36	36	-14	7.60	242
aSV > 0	vmPFC	-2	52	-8	3.61	481
	latOFC_L	-48	32	-8	6.04	203
vSV > 0	vmPFC	-4	58	-10	6.81	491
	postOFC_L	-36	24	-18	3.89	153

MNI coordinates (x, y, z) and T value corresponds to the local maxima peak of the cluster activations at SVFWE corrected P < 0.005 (cluster labels are from AAL atlas (Rolls et al., 2020)).

intraaudSV: contrast capturing responses elicited by changes in subjective value (SV) when choice options consisted of two auditory stimuli (*AudAud*). This contrast was calculated as *intraaudSV*>0 = *lpSV*>0 + *hpSV*>0, averaging responses to the low and high pitch auditory stimuli *(lpSV* and *hpSV*, respectively) against baseline.

intravisSV: contrast capturing responses elicited by changes in subjective value (SV) when choice options consisted of two visual stimuli (*VisVis*). This contrast was calculated as *intravisSV*>0 = *gSV*>0 + *rSV*>0, averaging responses to the green and red colors of visual stimuli (*gSV* and *rSV*, respectively) against baseline.

interaudvisSV: contrast capturing responses elicited by changes in subjective value (SV) when choice options consisted of one auditory and one visual stimulus (*AudVis*). This contrast was calculated as *interaudvisSV>0 = aSV>0 + vSV>0*, averaging responses to the auditory and visual stimuli (*aSV* and *vSV*, respectively) against baseline.

Together, these findings provide strong evidence that the valuation of stimuli from the auditory and visual sensory modalities is confined to segregated loci in OFC. Additionally, our results indicate that the representation of stimulus value is independent of sensory modality in vmPFC and that this region is involved in processing information related to the final choice across different tasks.

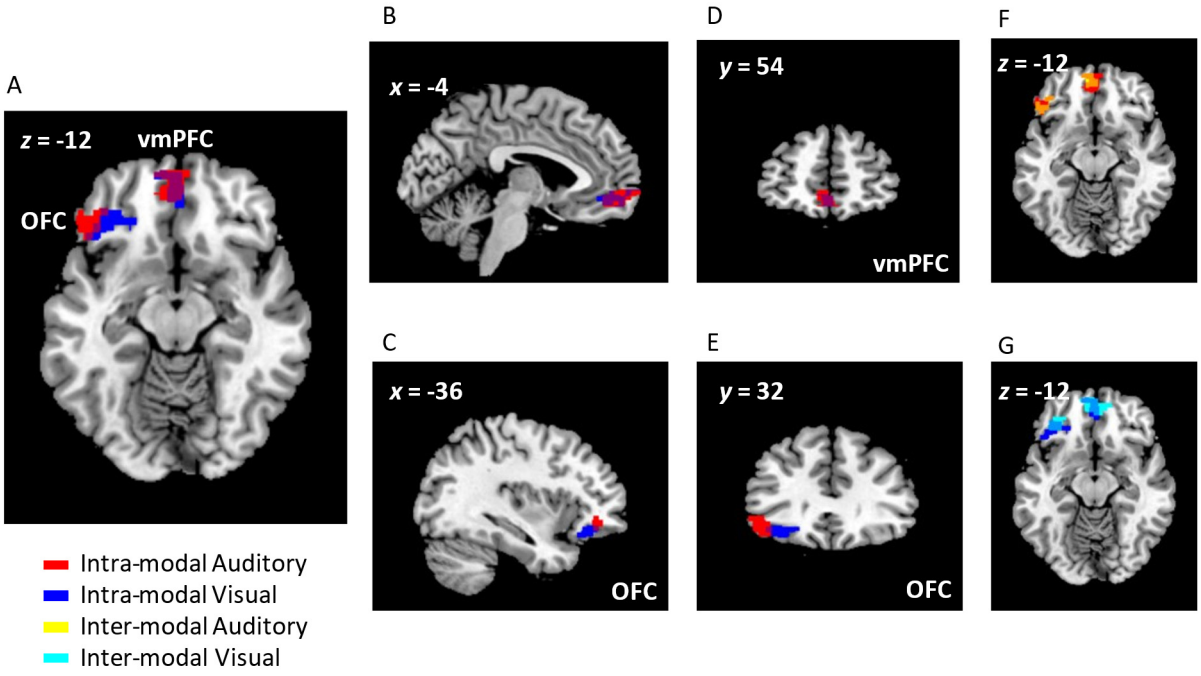

Figure 3. Stimulus value representations (SVR) across different sensory modalities. (A-E) In intra-modal conditions, segregated modality-specific SVR in OFC for auditory modality in left lateral OFC (red cluster) and for visual modality in left posterior OFC (blue cluster); and overlapping modality-general SVR in vmPFC for auditory and visual modalities (purple cluster due to overlap of red and blue). In inter-modal condition, the SVR for (F) auditory stimulus (corresponding to contrast $aSV > 0$) and (G) visual stimulus (corresponding to contrast $vSV > 0$), in valuation regions (OFC and vmPFC) were found to be overlapping with the SVRs of the same modality identified in intra-modal condition (as shown in (A), see also Figure S3). All cluster activations, shown here, are significant at SVFWE corrected $P < 0.005$.

Stimulus value representations outside of frontal valuation regions: Whole-brain analysis

Past studies have shown that reward value modulates the early sensory processing (Rutkowski and Weinberger, 2005; Shuler and Bear, 2006; Pleger et al., 2008; Serences, 2008; Goltstein et al., 2013). Thus, in order to identify regions exhibiting value modulations outside the valuation regions, specifically in auditory and visual sensory cortices, we performed a whole-brain analysis using the GLM described previously. For this purpose, we estimated the average effect of value across all conditions in the value task (*AudAud*, *VisVis* and *AudVis*), which revealed bilateral activations in the auditory and visual cortices (whole-brain FWE corrected $P < 0.05$, cluster size $k > 10$ voxels; **Table 2**, **Figure S5A** and **S5B**). Further, when estimating the value modulations for individual conditions separately (auditory, visual), we found modality-specific activations in respective sensory cortices only (see **Figure S5C-D**), whereas in intermodal condition both sensory cortices were activated (see **Figure S5E-F)**.

In addition to sensory cortices, we found significant value modulations in areas involved in processing of different aspects of value-related information, such as detecting the reward prediction errors (Caudate), formation of memories about past events (hippocampus), selection of action sets (SFGmed/ dmPFC) (Rushworth et al., 2004) and processing of

symbolic/linguistic information related to monetary value (Angular gyrus). Since the specific aim of the current study was to shed light on how modality-specific and modality-general valuation is coordinated across the frontal and sensory areas, we only included the whole-brain activations that were located in early visual or auditory areas in our subsequent effective connectivity analyses.

Table 2. Stimulus value representations outside of valuation regions

Region	*X*	*Y*	*Z*	*t*(19)	*K*
SFGmed	0	54	40	10.50	64
Caudate_L	-14	24	8	10.37	67
visSen_L	**-20**	**-90**	**2**	**9.75**	**81**
visSen_R	**24**	**-96**	**10**	**8.86**	**119**
Hippoc_L	-36	-38	-12	9.62	30
Hippoc_R	34	-22	-12	9.51	29
audSen_L	**-66**	**-30**	**-4**	**9.08**	**24**
audSen_R	**66**	**-12**	**-4**	**8.99**	**15**
Angular_L	-42	-56	26	8.62	55

MNI coordinates (x, y, z) and T value corresponds to the local maxima peak of the cluster activations at FWE corrected $P < 0.05$ (cluster labels are from AAL atlas (Rolls et al., 2020)). SFGmed – medial Superior Frontal Gyrus; Hippoc – Hippocampus; visSen – Visual Sensory Cortex; audSen – Auditory Sensory Cortex. Highlighted (in bold) activations were used as ROIs in the effective connectivity analysis.

Modality-specific effective connectivity between sensory and valuation areas

We next examined the effective connectivity (EC) of a network consisting of sensory and valuation regions that showed significant value-related modulations (i.e., 5 ROIs, see Material and Methods for details). The EC analysis (Friston et al., 2003) provides an estimation of the degree to which different connectivity patterns across this network contribute to the generation of modality-specific representations on the one hand and guide the final value-based choices on the other hand. The most probable connectivity pattern was captured by one out of 10 biologically plausible models (shown in **Figure 4A**) plus a null model, which was selected based on a Bayesian model comparison approach (Stephan et al., 2009).

We found that a model containing modulatory connections between the sensory cortices, modality-specific clusters in OFC and vmPFC was the most likely model in the value task (i.e. model 6, **Figure 4A-B**). Importantly, the winning model contained two distinct valuation sub-networks: an auditory sub-network comprising audSen, audOFC and vmPFC,

and a visual sub-network with visSen, visOFC and vmPFC as nodes (see **Table 2** for abbreviations of the ROIs). Moreover, the sensory cortices did not directly communicate with vmPFC (models 2, 4, 5, 7, 8 and 10) and we did not find evidence for the cross-modality of the connectivity between the sensory cortices and the value regions in OFC (models 1, 3, 4, 5, 9 and 10, containing a connection between visual cortex and auditory OFC or auditory cortex and visual OFC). These findings provide compelling evidence for the existence of modality-specific communication pathways which broadcast the value-related information across the brain.

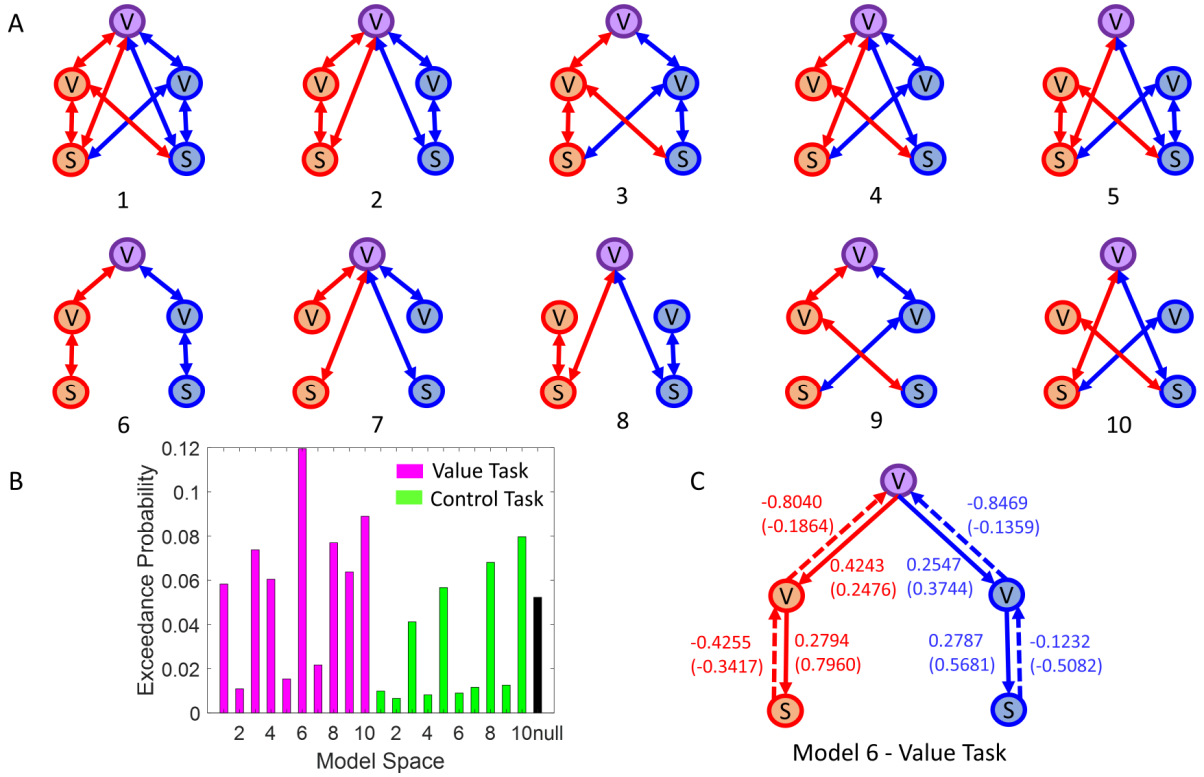

Figure 4. Modality-specific effective connectivity. (A) Model space consisting of 10 biologically plausible models (modulatory connections shown here) per task (value or control) plus a null model resulting into 21 potential models (Nodes: V – Valuation region, S – Sensory region, purple V – vmPFC, red V – Auditory OFC, blue V – Visual OFC, red S – Auditory Sensory Cortex, blue S – Visual Sensory Cortex). Modulatory connections in red exist during conditions when an auditory stimulus was selected and those in blue exist during conditions when a visual stimulus was selected in both intra- and inter-modal conditions. (B) Exceedance probabilities for the 21 potential connectivity models. Model 6 of value task is the most likely model. (C) Winning model shown with modulatory connection strength parameters for feedforward (dashed) and feedback (solid) connections during conditions when an auditory (red) or a visual (blue) stimulus either in intra-modal or inter-modal conditions (connection weights in brackets) was selected. All parameters significant at posterior probability of $P > 0.99$.

In order to understand how reward value modulates the communication of information across the brain areas, we next examined the strength of modulatory connections in the winning model. We found that during both intra-modal and inter-modal trials of the value task, all modality-specific connections were significantly modulated by value at a posterior probability $P > 0.99$ (**Figure 4C**). Interestingly, the connection strengths were negative for directed

feedforward connections from sensory ROIs to OFC ROIs to vmPFC, indicating inhibitory modulatory connections and positive for feedback connections from vmPFC to OFC ROIs to sensory ROIs, indicating excitatory modulatory connections (For parameters of intrinsic connections and driving inputs, refer to **Supplementary Information: Table S5-S6**)

Additionally, to test whether effective connectivity between any two nodes of the network was unidirectional, we estimated all possible unidirectional models for the winning model (i.e. model 6). As a modulatory connection between two nodes can exist in three possible ways: directed, reciprocal, bidirectional; the total number of all possible unidirectional models for the model 6 of the value task were 9, shown in **Figure 5A**. The base model for all these nine models was the same as for the bidirectional model space. Here again, the most likely model was the network with bidirectional connectivity between the nodes (see model exceedance probabilities, **Figure 5B**).

Together, the effective connectivity results showed that auditory and visual sensory cortices communicate with separate clusters in OFC, which contain modality-specific stimulus value representations (SVR) corresponding to each sensory modality. Further, the modality-specific SVR in OFC were linked with modality-general SVR in vmPFC to guide the final value-based choices.

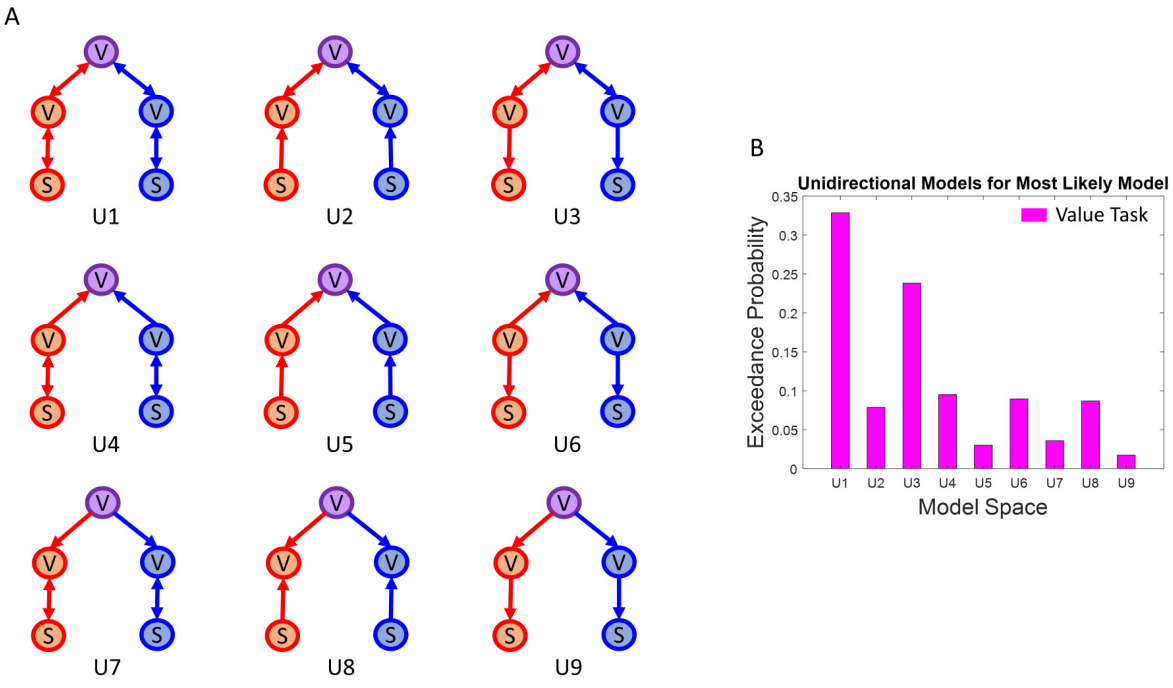

Figure 5. Uni- and bi-directional variants of the winning model in the value task. (A) Model space consisting of all possible unidirectional models (details in text; also refer to the legend of Figure 4A for information on nodes and modulatory connections). (B) Exceedance probabilities for the model space in (A). Model U1, i.e., a model containing bidirectional connections, is the most likely model.

Discussion

In order to generate specific predictive signals for adaptive goal-directed choices, the brain must encode information about the sensory modality of reward predicting stimuli as well as the most recent value associations with the stimuli. Moreover, to be able to compare and choose between stimuli having fundamentally distinct sensory features, general value representations are equally important. Here, we used stimuli from auditory and visual sensory modalities as reward-predicting cues in a value-based decision-making task with a dynamic foraging paradigm, enabling us to dissociate auditory and visual value representations using univariate fMRI analysis and examine the underlying neural mechanisms of goal-directed choice using effective connectivity analyses. We found that modality-specific value representations in OFC played a central role in generation of a modality-specific valuation network involving the value representations encoded in the sensory cortices and the modality-general value representations in vmPFC.

We found trial-by-trial value representations of auditory and visual sensory modalities to be present in segregated non-overlapping lateral and posterior regions of OFC, respectively. Recent studies have proposed OFC as a key neural substrate for supporting the formation of a "cognitive-map" of the current task space (Wilson et al., 2014; Stalnaker et al., 2015). Such a cognitive map of task space is fundamental for goal-directed behaviour by keeping track of all possible relevant states of the environment, required for generating specific predictive information about upcoming decisions such as those related to their specific sensory features, especially when a task involves reversal learning (Tsuchida et al., 2010; Stalnaker et al., 2015) or the states change in a way that require devaluation of previously valuable options (Pickens et al., 2003). We employed a behavioural paradigm which varied value independently of the modality of the reward options. The dynamic reward structure of the value task required participants to track the reward history of each sensory modality on a trial-by-trial basis, as these could change over time. Using a task that required rapid updating of value associations of auditory and visual stimuli, which draws on the role of OFC in representing a map of task space, we found segregated value representations for each sensory modality. Interestingly, this segregation existed not only when both options were from different sensory domains competing against each other but also when the options were from the same modality domain. Furthermore, we verified using the control task that the segregation in modality-specific representations does not exist due to differences in sensory processing mechanism underlying the auditory and visual sensory modalities. Thus, our findings show for the first time that

dedicated neuronal populations exist in OFC for individual sensory modalities that encode value separately to reflect updates in value associations of a particular modality and generate specific predictive signals. As such, the present study extends our understanding of OFC's role in goal-directed behaviour to include the implementation of a modality-specific cognitive map of the task space.

In contrast to the modality-specific value representations found in OFC, we found modality-general value representations in vmPFC. Here, the auditory value representations were found overlapping with the visual value representations. This is in line with previous studies highlighting vmPFC as a common currency coding hub for distinct reward categories such as money, food, social rewards (Hare et al., 2008, 2010, 2011; Chib et al., 2009; Lebreton et al., 2009; Rolls et al., 2010; Smith et al., 2010; Levy and Glimcher, 2011; Lin et al., 2012; McNamee et al., 2013). However, previous work has also highlighted the idea that subdivisions from general to specific valuation exist in vmPFC in the posterior-to-anterior direction (Kringelbach and Rolls, 2004; Sescousse et al., 2010; Smith et al., 2010; Clithero and Rangel, 2013; McNamee et al., 2013), where anterior vmPFC represents values of distinct reward categories in a general manner and posterior vmPFC in a specific manner. The loci of overlapping activations, which we found in this study, were in anterior vmPFC and thus in line with the role of the anterior vmPFC in common currency coding of value. However, we did not find any modality-specific value representations in vmPFC, which may either be due to OFC being exclusively responsible for implementing modality-specificity in a task such as ours or be related to the specific type of reward category, i.e., monetary rewards, that we employed in our task (McNamee et al., 2013). Future studies will be needed to reveal whether posterior vmPFC will undertake a role in representing the identity of rewards of different sensory modalities when they are associated with other categories of reward other than monetary value. Interestingly, we also found that vmPFC activations, albeit at a weak level, were also present in the control task where no continuous and gradual value-related information processing was needed. This finding is in line with recent theoretical frameworks suggesting a general role of vmPFC in computation of choice rather than valuation and further calls for revisiting the common currency models of valuation (Klein-Flügge et al., 2022).

Whereas the majority of previous studies have underscored a common currency coding of reward value, few recent studies have provided evidence for the identity-specific representations of value (Klein-Flügge et al., 2013; McNamee et al., 2013; Howard et al., 2015; Howard and Kahnt, 2017) and affect/valence (Čeko et al., 2022). In fact, identity-specific value representations in previous studies could only be identified when highly sensitive data

acquisition or analysis of fMRI data were undertaken, for instance BOLD adaptation suppression (Klein-Flügge et al., 2013) or multivariate analysis across a range of different types of rewards in addition to and apart from monetary rewards (McNamee et al., 2013). We are only aware of one recent imaging study with humans (Shuster and Levy, 2018), where modality-specificity of valuation across auditory and visual domains was examined. Using a risk-evaluation task where lotteries were presented either visually or aurally they found that the anterior portion of vmPFC represents value irrespective of the sensory-modality, whereas no evidence for modality-specificity was found beyond sensory areas. However, in their study (Shuster and Levy, 2018) modality-specific value information was rendered redundant per design of the experiment, because the decisions involved comparison of explicit lottery options associated with different amounts of monetary risk and a safe option, in which lotteries' monetary value were either presented visually or aurally. Since both visual and auditory stimuli could be translated to the same abstract numeric value and no dynamic change in options sensory identity occurred, valuation process could entirely occur without tracking the sensory modality. In contrast, in the present study, we used a design in which trial-by-trial updating of computed values of specific identity of sensory stimuli was necessary, thus allowing us to tap into the intricate role of OFC in modality-specific updating of value. Another important aspect of our approach was to account for the covariation of sensory features and reward value (Howard and Kahnt, 2021) in determining neuronal responses. This was done by employing a control task that was identical to the value task in terms of sensory requirements and final choice but differed in whether updating of computed reward value of each sensory modality was necessary or not. Together, a dynamic reward structure and the comparison against a task with different dimension of decision variable allowed us to unravel the co-existence of modality-specific and modality-general representations in the frontal cortex.

Apart from the frontal cortex, we found value modulations in sensory cortices, which provide evidence that representations of value are not restricted only to higher cognitive areas, as has been shown before (Serences, 2008). The value representations in sensory cortices were largely modality-specific, which means that individual sensory cortices represented the value of stimuli presented in their own sensory domain, a finding that is in line with previous studies on representation of value (Shuster and Levy, 2018) and valence (Čeko et al., 2022). These findings raised interesting questions regarding whether and how a communication of value-related information exists between sensory cortices and valuation regions. Interestingly, we found that the auditory and visual sensory cortices were bi-directionally connected to the lateral and posterior OFC (corresponding to auditory and visual value representations), respectively,

in a modality-specific manner. Specifically, the modality-specific effective connectivity results revealed a high degree of selectivity: in a trial when planning to choose auditory reward stimulus, there was a significant connectivity from the auditory sensory cortex to lateral OFC for that trial and not otherwise. A similar modality-specific significant connectivity existed from visual sensory cortex to posterior OFC for choosing visual reward stimulus. This finding is in line with a previous work showing connectivity between OFC and piriform cortex (relevant in case of odour stimuli) for the formation of identity-specific value representations in OFC (Howard et al., 2015). Moreover, past studies have shown that lateral and posterior regions of OFC receive direct afferent inputs from auditory and visual sensory cortices (Barbas, 1993; Carmichael and Price, 1995), providing neuroanatomical support for our findings.

The connectivity between sensory cortices and modality-specific representations in OFC reveals an underlying mechanism by which modality-specific sensory features of a reward option are extracted from the respective sensory cortex, and then further processed in OFC (and not in vmPFC) along with value information to support formation of modality-specific value codes. The sign of this connectivity provides additional information regarding how modality-specific valuation is implemented. We found that feedforward connectivity in modality-specific networks was predominantly inhibitory. The inhibitory feedforward connectivity indicates that when choices involve a specific sensory modality, early sensory areas send inhibitory signals to OFC. Feedforward inhibition has been suggested as a key mechanism in imposing temporal structure to neuronal responses (Womelsdorf and Everling, 2015) and expanding their dynamic range of activity (Pouille et al., 2009). This mechanism highlights the role of OFC in the formation of an integrated memory trace of the sensory and value information over time, rather than encoding the exact sensory features of stimuli at each instance. This finding is supported by a notion of OFC as a cognitive map that stores task space over time, as proposed before (Wilson et al., 2014; Stalnaker et al., 2015). Additionally, connectivity results showed that the value modulations in sensory cortices were driven by top-down feedback signals generated in respective valuation regions in OFC. This is in line with previous work showing that biasing signals generated from frontal and parietal areas modulate spatially selective visual areas (Serences, 2008). In fact, recent studies have provided robust causal evidence for the role of lateral OFC in value-driven guidance of information processing in sensory cortices (Banerjee et al., 2020). Our finding of the presence of excitatory feedback connectivity between the modality-specific representations in lateral and posterior OFC and auditory and visual cortices, provides strong support for the causal role of top-down valuation

902 signals in shaping sensory perception during decision-making, through enhancing the sensory
903 information that is most relevant for the current choice.
904 Further, we found that specific value representation in OFC were linked to general value
905 representations in vmPFC. Specifically, we showed that when planning to select an auditory
906 reward option, there was a change in the connectivity between the auditory value
907 representations in OFC and modality-general representations in vmPFC, with a similar pattern
908 found for the selection of the visual reward options. This result highlights the underlying
909 mechanism whereby value representations in OFC provide input to the vmPFC to support the
910 formation of general value representations needed for the comparison of options from distinct
911 domains and deriving the final choice. These finding are in line with the role of OFC in
912 providing fine-tuned value information that can drive the choice (Setogawa et al., 2019) and
913 the role of vmPFC in the final comparison and computation of choice that guides actions (Hare
914 et al., 2011). Importantly, the modality-specific connectivity between OFC and vmPFC is in
915 line with a previous work showing that sensory-specific satiety-related changes in connectivity
916 between OFC and vmPFC predicted choices in a devaluation task (Howard and Kahnt, 2017).
917 Together, these results show how common currency coding of value integrates identity-specific
918 information about reward options in a dynamic environment to guide choices.
919 Understanding whether valuation signals in frontal cortex contain information about the
920 sensory modality of rewarded stimuli has a number of important theoretical and clinical
921 implications that go beyond the specialized field of neuroeconomics and value-based decision
922 making (Rangel et al., 2008; Levy and Glimcher, 2012; Padoa-Schioppa and Schoenbaum,
923 2015). We show that value-based choices involving reward options with distinct sensory
924 features are supported by bi-directional connectivity between the sensory areas and the
925 modality-specific representations in OFC. Although the top-down modulation of perception
926 through interactions between frontal and sensory areas has been the basic tenet of a number of
927 influential theoretical frameworks (Desimone and Duncan, 1995; Corbetta and Shulman, 2002;
928 Friston, 2005; Gardner and Schoenbaum, 2021), the importance of modality-specific
929 representations of reward value in frontal areas that could provide a biologically plausible
930 implementation of these putative interactions has been largely ignored. Therefore, our study
931 provides novel insight for future computational work on how top-down signals can be
932 selectively routed to impact on sensory processing. In doing so, it is important to note that the
933 modality representations that we found may adapt and reorganize under different contexts
934 rather than being hardwired and fixed in the brain. In fact, outcome-related adaptation in the
935 representation of value can occur during the same task (Rich and Wallis, 2016), which provides

a flexible mechanism for reorganizing neuronal codes of value based on the context. Future studies will be needed to examine whether and to what extent the modality-specific coding of value can adapt to the specific features of a task. From a clinical perspective, our results suggest that localized lesions to OFC may be associated with specialized impairments of value-based decisions in visual or auditory domains, an interesting possibility that can be further investigated by future studies. Additionally, our findings may allow a better understanding of pathological states such hallucinations (Frith, 1996; Rolls et al., 2008) where illusory percepts arise in the absence of external stimuli (Powers et al., 2016), likely due to the aberrations in communication pathways between the frontal and sensory areas (Allen et al., 2008). More generally, the present study, together with previous efforts in understanding how value-related information is communicated between the frontal and sensory areas (Howard et al., 2015; Howard and Kahnt, 2018; Banerjee et al., 2020), provide instrumental insights regarding how perceptual and cognitive processes are coordinated in the brain.

In summary, our results provide evidence for the co-existence of modality-specific and modality-general codes in OFC and vmPFC, respectively, pointing to the specialized functions of these two valuation areas. A general value signal would facilitate the comparison between distinct rewards (Levy and Glimcher, 2011, 2012) and the transformation of stimulus values into motor commands (Hare et al., 2011). On the contrary, modality-specific value encoding associated to respective sensory cortical representations would support goal-directed adaptive behaviour by generating specific predictive signals about impending goals (Stalnaker et al., 2014; Wilson et al., 2014; Nogueira et al., 2017), such as when planning to choose auditory or visual reward stimuli. We further show how the communications between sensory areas and modality-specific representations of reward value in OFC play a central role in supporting value-based decisions in a multimodal dynamic environment.

Acknowledgements

We thank Tabea Hildebrand and Jana Znaniewitz for their help with the data collection. This work was supported by an ERC Starting Grant (no: 716846) to AP.

Authors' contributions

SD and AP conceptualized the project and designed the main task. SD, IK and AP designed the control task. SD, JEA, and AP conducted the experiments. JEA preprocessed the data. SD

and AP analyzed the data. SD and AP interpreted the results and wrote the first draft of the manuscript. All authors revised the manuscript. AP acquired funding.

References

Allen P, Larøi F, McGuire PK, Aleman A (2008) The hallucinating brain: A review of structural and functional neuroimaging studies of hallucinations. Neurosci Biobehav Rev 32:175–191.

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. In: Neuropharmacology, pp 407–419.

Banerjee A, Parente G, Teutsch J, Lewis C, Voigt FF, Helmchen F (2020) Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. Nature 585:245–250.

Barbas H (1988) Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey. J Comp Neurol 276:313–342.

Barbas H (1993) Organization of cortical afferent input to orbitofrontal areas in the rhesus monkey. Neuroscience 56:841–864.

Baxter MG, Parker A, Lindner CCC, Izquierdo AD, Murray EA (2000) Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J Neurosci 20:4311–4319.

Berridge KC, Kringelbach ML (2015) Pleasure Systems in the Brain. Neuron 86:646–664.

Burks JD, Conner AK, Bonney PA, Glenn CA, Baker CM, Boettcher LB, Briggs RG, O'Donoghue DL, Wu DH, Sughrue ME (2018) Anatomy and white matter connections of the orbitofrontal gyrus. J Neurosurg 128:1865–1872.

Cai X, Padoa-Schioppa C (2014) Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. Neuron 81:1140–1151.

Carmichael ST, Price JL (1995) Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 363:642–664.

Carmichael ST, Price JL (1996) Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 371:179–207.

Čeko M, Kragel PA, Woo CW, López-Solà M, Wager TD (2022) Common and stimulus-type-specific brain representations of negative affect. Nat Neurosci 25:760–770.

Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. J Neurosci 29:12315–12320.

Clithero JA, Rangel A (2013) Informatic parcellation of the network involved in the computation of subjective value. Soc Cogn Affect Neurosci 9:1289–1302.

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215.

Corrado GS, Sugrue LP, Sebastian Seung H, Newsome WT (2005) Linear-Nonlinear-Poisson Models of Primate Choice Dynamics. J Exp Anal Behav 84:581–617.

1000 Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222.

1001 Eryilmaz H, Rodriguez-Thompson A, Tanner AS, Giegold M, Huntington FC, Roffman JL (2017) Neural determinants of
1002 human goal-directed vs. habitual action control and their relation to trait motivation. Sci Rep 7.

1003 Friston K (2005) A theory of cortical responses. Philos Trans R Soc B Biol Sci 360:815–836.

1004 Friston KJ (2011) Functional and Effective Connectivity: A Review. Brain Connect 1:13–36.

1005 Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. Neuroimage 19:1273–1302.

1006 Friston KJ, Li B, Daunizeau J, Stephan KE (2011) Network discovery with DCM. Neuroimage 56:1202–1221.

1007 Frith C (1996) The role of the prefrontal cortex in self-consciousness: the case of auditory hallucinations. Philos Trans R Soc
1008 London Ser B Biol Sci 351:1505–1512.

1009 Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative
1010 learning. J Neurosci 19:6610–6614.

1011 Gardner MPH, Schoenbaum G (2021) The orbitofrontal cartographer. Behav Neurosci 135:267–276.

1012 Goltstein PM, Coffey EBJ, Roelfsema PR, Pennartz CMA (2013) In vivo two-photon Ca2+ imaging reveals selective reward
1013 effects on stimulus-specific assemblies in mouse visual cortex. J Neurosci 33:11540–11555.

1014 Hallett M (1998) Human Brain Function.

1015 Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference
1016 during decision making in humans. J Neurosci 26:8360–8367.

1017 Hare TA, Camerer CF, Knoepfle DT, Rangel A (2010) Value computations in ventral medial prefrontal cortex during
1018 charitable decision making incorporate input from regions involved in social cognition. J Neurosci 30:583–590.

1019 Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the
1020 striatum in the computation of goal values and prediction errors. J Neurosci 28:5623–5630.

1021 Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value signals into motor
1022 commands during simple choice. Proc Natl Acad Sci U S A 108:18120–18125.

1023 Herrnstein RJ, Baum WM (1970) Journal of the Experimental Analysis of Behavior on the Law of Effect'. 2:243–266.

1024 Hiser J, Koenigs M (2018) The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making, Social
1025 Cognition, and Psychopathology. Biol Psychiatry 83:638–647.

1026 Howard JD, Gottfried JA, Tobler PN, Kahnt T (2015) Identity-specific coding of future rewards in the human orbitofrontal
1027 cortex. Proc Natl Acad Sci U S A 112:5195–5200.

1028 Howard JD, Kahnt T (2017) Identity-specific reward representations in orbitofrontal cortex are modulated by selective
1029 devaluation. J Neurosci 37:2627–2638.

1030 Howard JD, Kahnt T (2018) Identity prediction errors in the human midbrain update reward-identity expectations in the
1031 orbitofrontal cortex. Nat Commun 9:1–11.

1032 Howard JD, Kahnt T (2021) To be specific: The role of orbitofrontal cortex in signaling reward identity. Behav Neurosci
1033 135:210–217.

1034 Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided
1035 by both reward value and reward contingency. J Neurosci 24:7540–7548.

1036 Katahira K (2015) The relation between reinforcement learning parameters and the influence of reinforcement history on
1037 choice behavior. J Math Psychol 66:59–69.

1038 Kennerley SW, Dahmubed AF, Lara AH, Wallis JD (2009) Neurons in the frontal lobe encode the value of multiple decision
1039 variables. J Cogn Neurosci 21:1162–1178.

1040 Klein-Flügge MC, Barron HC, Brodersen KH, Dolan RJ, John Behrens TE (2013) Segregated encoding of reward-identity and
1041 stimulus-reward associations in human orbitofrontal cortex. J Neurosci 33:3202–3211.

1042 Klein-Flügge MC, Bongioanni A, Rushworth MFS (2022) Medial and orbital frontal cortex in decision-making and flexible
1043 behavior. Neuron 110.

1044 Kötter R, Stephan KE (2003) Network participation indices: Characterizing component roles for information processing in
1045 neural networks. Neural Networks 16:1261–1275.

1046 Kringelbach ML, Rolls ET (2004) The functional neuroanatomy of the human orbitofrontal cortex: Evidence from
1047 neuroimaging and neuropsychology. Prog Neurobiol 72:341–372.

1048 Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An Automatic Valuation System in the Human Brain:
1049 Evidence from Functional Neuroimaging. Neuron 64:431–439.

1050 Levy DJ, Glimcher PW (2011) Comparing apples and oranges: Using reward-specific and reward-general subjective value
1051 representation in the brain. J Neurosci 31:14693–14707.

1052 Levy DJ, Glimcher PW (2012) The root of all value: A neural common currency for choice. Curr Opin Neurobiol 22:1027–
1053 1038.

1054 Lin A, Adolphs R, Rangel A (2012) Social and monetary reward learning engage overlapping neural substrates. Soc Cogn
1055 Affect Neurosci 7:274–281.

1056 Lopez-Persem A, Roumazeilles L, Folloni D, Marche K, Fouragnan EF, Khalighinejad N, Rushworth MFS, Sallet J (2020)
1057 Differential functional connectivity underlying asymmetric reward-related activity in human and nonhuman primates.
1058 Proc Natl Acad Sci U S A 117:28452–28462.

1059 Mannella F, Mirolli M, Baldassarre G (2016) Goal-directed behavior and instrumental devaluation: A neural system-level
1060 computational model. Front Behav Neurosci 10.

1061 Martínez-Molina N, Mas-Herrero E, Rodríguez-Fornells A, Zatorre RJ, Marco-Pallarés J (2019) White matter microstructure
1062 reflects individual differences in music reward sensitivity. J Neurosci 39:5018–5027.

1063 McNamee D, Rangel A, O'Doherty JP (2013) Category-dependent and category-independent goal-value codes in human
1064 ventromedial prefrontal cortex. Nat Neurosci 16:479–485.

1065 Montague PR, Berns GS (2002) Neural economics and the biological substrates of valuation. Neuron 36:265–284.

1066 Nogueira R, Abolafia JM, Drugowitsch J, Balaguer-Ballester E, Sanchez-Vives M V., Moreno-Bote R (2017) Lateral
1067 orbitofrontal cortex anticipates choices and integrates prior with current information. Nat Commun 8.

1068 Noonan MP, Mars RB, Rushworth MFS (2011) Distinct roles of three frontal cortical areas in reward-guided behavior. J
1069 Neurosci 31:14399–14412.

1070 Noonan MP, Walton ME, Behrens TEJ, Sallet J, Buckley MJ, Rushworth MFS (2010) Separate value comparison and learning
1071 mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci U S A 107:20547–20552.

1072 O'Doherty JP (2014) The problem with value. Neurosci Biobehav Rev 43:259–268.

1073 O'Doherty JP, Cockburn J, Pauli WM (2017) Learning, Reward, and Decision Making. Annu Rev Psychol 68:73–100.

1074 Padoa-Schioppa C (2011) Neurobiology of Economic Choice: A Good-Based Model. Annu Rev Neurosci.

1075 Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. Nature 441:223–226.

1076 Padoa-Schioppa C, Assad JA (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes
1077 of menu. Nat Neurosci 11:95–102.

1078 Padoa-Schioppa C, Schoenbaum G (2015) Dialogue on economic choice, learning theory, and neuronal representations. Curr
1079 Opin Behav Sci 5:16–23.

1080 Pickens CLL, Saddoris MPP, Setlow B, Gallagher M, Holland PCC, Schoenbaum G (2003) Different Roles for Orbitofrontal
1081 Cortex and Basolateral Amygdala in a Reinforcer Devaluation Task. J Neurosci 23:11078–11084.

1082 Pleger B, Blankenburg F, Ruff CC, Driver J, Dolan RJ (2008) Reward facilitates tactile judgments and modulates
1083 hemodynamic responses in human primary somatosensory cortex. J Neurosci 28:8161–8168.

1084 Poline JB, Holmes AP, Worsley KJ, Friston KJ (1997) Making statistical inferences. Hum Brain Funct:85–106.

1085 Pouille F, Marin-Burgin A, Adesnik H, Atallah B V., Scanziani M (2009) Input normalization by global feedforward inhibition
1086 expands cortical dynamic range. Nat Neurosci 12:1577–1585.

1087 Powers AR, Kelley M, Corlett PR (2016) Hallucinations as Top-Down Effects on Perception. Biol Psychiatry Cogn Neurosci
1088 Neuroimaging 1:393–400.

1089 Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. Nat
1090 Rev Neurosci 9:545–556.

1091 Rich EL, Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19:973–980.

1092 Rolls ET (2000) The orbitofrontal cortex and reward. Cereb Cortex 10:284–294.

1093 Rolls ET, Grabenhorst F, Parris BA (2010) Neural systems underlying decisions about affective odors. J Cogn Neurosci
1094 22:1069–1082.

1095 Rolls ET, Huang CC, Lin CP, Feng J, Joliot M (2020) Automated anatomical labelling atlas 3. Neuroimage 206:116189.

1096 Rolls ET, Joliot M, Tzourio-Mazoyer N (2015) Implementation of a new parcellation of the orbitofrontal cortex in the
1097 automated anatomical labeling atlas. Neuroimage 122:1–5.

1098 Rolls ET, Loh M, Deco G, Winterer G (2008) Computational models of schizophrenia and dopamine modulation in the
1099 prefrontal cortex. Nat Rev Neurosci 9:696–709.

1100 Rudebeck PH, Murray EA (2011) Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on
1101 reward-guided behavior. J Neurosci 31:10569–10578.

1102 Rudebeck PH, Murray EA (2014) The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific
1103 behavioral outcomes. Urology 84:1143–1156.

1104 Rudebeck PH, Saunders RC, Lundgren DA, Murray EA (2017) Specialized Representations of Value in the Orbital and
1105 Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes. Neuron 95:1208-1220.e5.

1106 Rushworth MFS, Walton ME, Kennerley SW, Bannerman DM (2004) Action sets and decisions in the medial frontal cortex.
1107 Trends Cogn Sci 8:410–417.

1108 Rutkowski RG, Weinberger NM (2005) Encoding of learned importance of sound by magnitude of representational area in
1109 primary auditory cortex. Proc Natl Acad Sci U S A 102:13664–13669.

1110 Serences JT (2008) Value-based modulations in human visual cortex. Neuron 60:1169–1181.

1111 Sescousse G, Redouté J, Dreher JC (2010) The architecture of reward value coding in the human orbitofrontal cortex. J
1112 Neurosci 30:13095–13104.

1113 Setogawa T, Mizuhiki T, Matsumoto N, Akizawa F, Kuboki R, Richmond BJ, Shidara M (2019) Neurons in the monkey
1114 orbitofrontal cortex mediate reward value computation and decision-making. Commun Biol 2.

1115 Shuler MG, Bear MF (2006) Reward timing in the primary visual cortex. Science 311:1606–1609.

1116 Shuster A, Levy DJ (2018) Common sense in choice: The effect of sensory modality on neural value representations. eNeuro
1117 5.

1118 Smith D V., Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior
1119 ventromedial prefrontal cortex. J Neurosci 30:2490–2495.

1120 Stalnaker TA, Cooch NK, McDannald MA, Liu TL, Wied H, Schoenbaum G (2014) Orbitofrontal neurons infer the value and
1121 identity of predicted outcomes. Nat Commun 5.

1122 Stalnaker TA, Cooch NK, Schoenbaum G (2015) What the orbitofrontal cortex does not do. Nat Neurosci 18:620–627.

1123 Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. Neuroimage
1124 46:1004–1017.

1125 Suzuki S, Cross L, O'Doherty JP (2017) Elucidating the underlying components of food valuation in the human orbitofrontal
1126 cortex. Nat Neurosci 20:1780–1786.

1127 Tsuchida A, Doll BB, Fellows LK (2010) Beyond reversal: A critical role for human orbitofrontal cortex in flexible learning
1128 from probabilistic feedback. J Neurosci 30:16868–16875.

1129 Valentin V V., Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human
1130 brain. J Neurosci 27:4019–4026.

1131 Vandekerckhove J, Matzke D, Wagenmakers E-J (2015) Model Comparison and the Principle of Parsimony (Busemeyer JR,
1132 Wang Z, Townsend JT, Eidels A, eds). Oxford University Press.

1133 Wallis JD (2012) Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat Neurosci 15:13–19.

1134 Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a
1135 reward preference task. Eur J Neurosci 18:2069–2081.

1136 Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron
1137 81:267–279.

1138 Womelsdorf T, Everling S (2015) Long-Range Attention Networks: Circuit Motifs Underlying Endogenously Controlled

Stimulus Selection. Trends Neurosci 38:682–700.

Wunderlich K, Dayan P, Dolan RJ (2012) Mapping value based planning and extensively trained choice in the human brain. Nat Neurosci 15:786–791.

Zeki S, Shipp S (1988) The functional logic of cortical connections. Nature 335:311–317.

Supplementary Information to "Modality-specific and modality-general representations of reward value in frontal cortex"

Dang et al.

Supplementary Figures

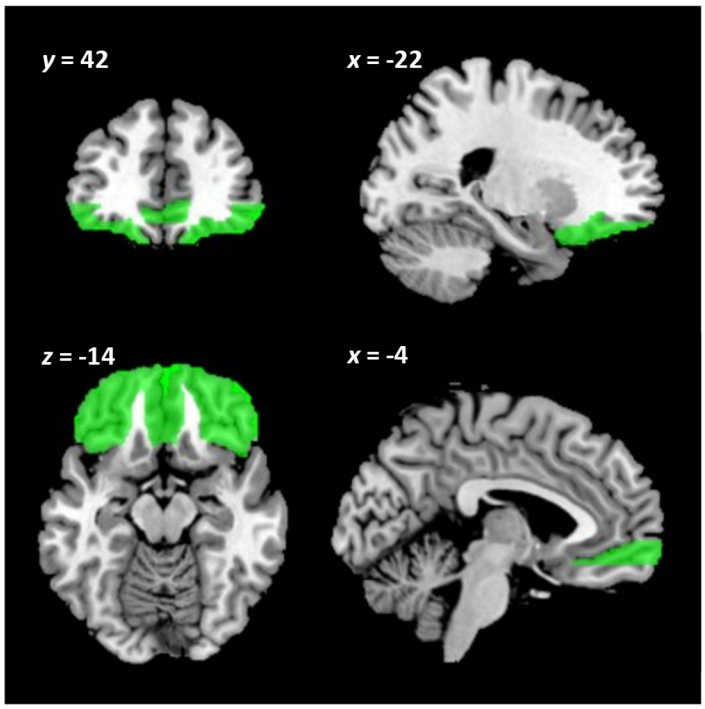

Figure S1. Anatomical definition of frontal valuation areas. The search volume used for multiple comparisons correction consisted of anatomical parcellations of the orbital surface of frontal gyrus as defined in automated anatomical labelling (AAL) atlas (Rolls et al., 2015, 2020). The search volume, comprised the anatomical parcellations of orbital surface in the following format, ROI name (abbreviation): Superior frontal gyrus - medial orbital (PFCventmed); Medial orbital gyrus (OFCmed); Anterior orbital gyrus (OFCant); Posterior orbital gyrus (OFCpost); Lateral orbital gyrus (OFClat), for the detailed description of these areas see Table 2 in Rolls et al. 2015, 2020 (Rolls et al., 2015, 2020).

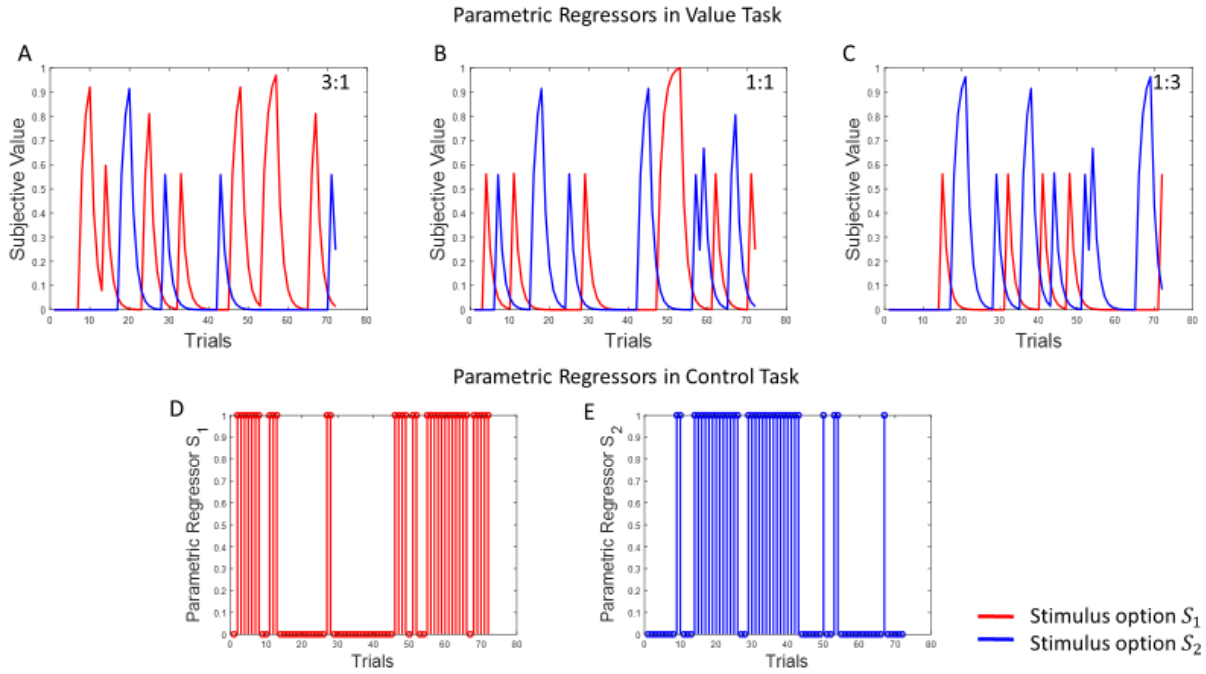

Figure S2. Illustration of the time course of parametric regressors in the value and control tasks. Parametric regressors used for the fMRI analysis are shown for a single participant: (A), (B), (C) For the value task, subjective values (SVs) of each option (S_1, S_2) across trials in a block with reward ratios of 3:1, 1:1, 1:3, respectively, are shown. SVs were calculated based on the computational modeling of behavioural data (see Material and Methods in the main text). (D), (E) For the control task, where instructions where passively followed across trials in a block, a weight of 0 or 1 was assigned to each option. The weights assigned to S_1 and S_2 in the control task were determined based on the schema shown in **Table S1**.

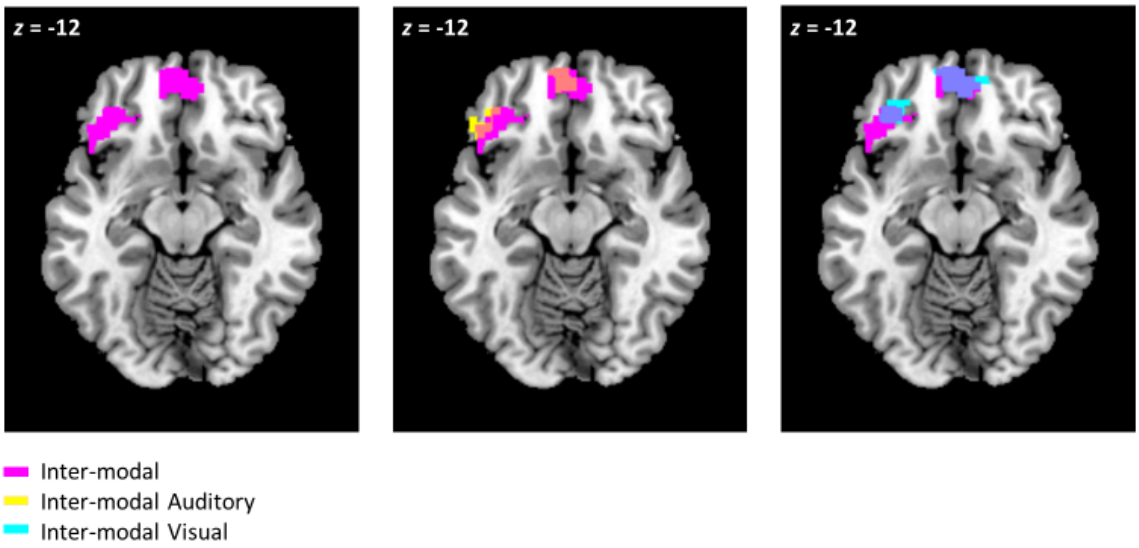

Figure S3. Stimulus value representations (SVR) in inter-modal condition. This figure should be compared to Figure 3 in the main text. Here we illustrate the SVRs in the frontal valuation areas (search volume shown in Figure S1) only for the inter-modal condition. The activations in magenta correspond to the contrast *interaudvisSV>0*, the activations in yellow correspond to the contrast *aSV>0*, and the activations in cyan correspond to the contrast *vSV>0* (for definition of contrasts see Table 1 in the main text). All cluster activations shown here are significant at SVFWE corrected P < 0.005.

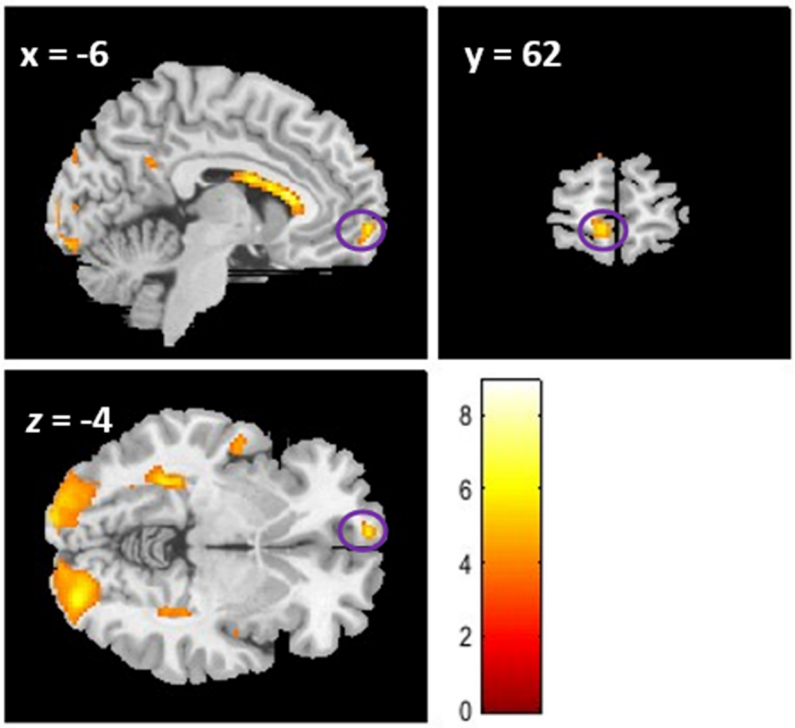

Figure S4. Activation in vmPFC for the control task shown across all conditions at whole-brain uncorrected level of P < 0.001.

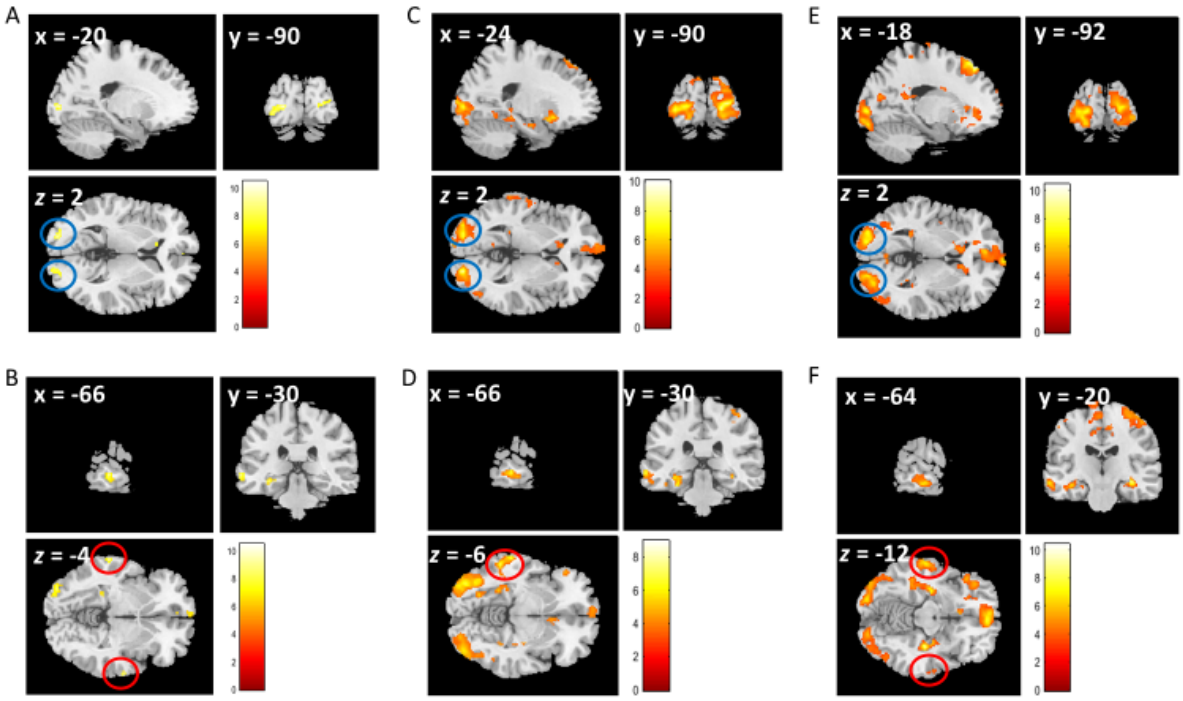

Figure S5. Stimulus value representations in sensory cortices: (A) Visual sensory cortex. (B) Auditory sensory cortex. All cluster activations shown in (A) and (B) were significant at whole-brain FWE corrected $P < 0.05$ and are estimated across all conditions (*AudAud, VisVis and AudVis*). (C) When the value modulations were inspected for individual conditions separately (whole-brain uncorrected level of $P < 0.001$), examination of the contrast *intravisSV* > 0 revealed activations in the visual sensory cortex (cluster peaks at (-24, -90, 2) and (18, -94, 8)) but no activation in the auditory sensory cortex were found. (D) Similarly, for the contrast *intraaudSV* > 0, we found activations in the auditory sensory cortex (cluster peak at (-66, -30, -6)) but no activation in the visual sensory cortex. This result indicates that each sensory cortex was specifically activated when the value of a stimulus from its specific modality was processed. However, for this contrast, we also found activations in higher visual areas (occipitotemporal cortex) with cluster peaks at (-36, -66, -8) and (42, -76, -12) that were distinct from activations found for the contrast *intravisSV* > 0 which were in early visual areas in the occipital cortex (anatomical definitions are based on https://neurosynth.org/). (E) For the contrast *interaudvisSV* > 0, we found activations in both the visual sensory cortex (cluster peaks at (-18, -92, 2) and (22, -90, 6)) and the auditory sensory cortex (cluster peaks at (-64, -20, -12) and (66, -10, -4)), as in audio-visual condition trial-by-trial subjective values are updated individually for both auditory and visual options (whole-brain uncorrected level of $P < 0.001$). In all figures, crosshairs are placed at the left hemisphere cluster peak.

Supplementary Text and Tables

Definition of parametric regressors for the control task

Similar to the value task, two regressors were modelled at the onset of stimuli options in the control task: one unmodulated regressor representing the modality-wise trial identity and two parametrically modulated regressors for each of the two choice options (S_1 and S_2). To define the parametric regressors, we assigned a weight of either 1 or 0 to each option according to the schema shown in **Table S1**.

Table S1. Definition of parametric regressors for the control task*

Choice on trial *t*-1	Instruction from trial *t*-1	Weight assigned on trial *t*
S_1	Not followed	S_1 - 0, S_2 - 0
S_2	Not followed	S_1 - 0, S_2 - 0
S_1	Followed; Instruction was to keep	S_1 - 1, S_2 - 0
S_1	Followed; Instruction was to switch	S_1 - 0, S_2 - 1
S_2	Followed; Instruction was to keep	S_1 - 0, S_2 - 1
S_2	Followed; Instruction was to switch	S_1 - 1, S_2 - 0

*When instruction from the previous trial (*t*-1) was not followed, a weight of 0 was assigned to both options S_1 and S_2. When the instruction from the previous trial *t*-1 was correctly followed, the option that corresponded to the correct instructed choice in trial *t* received a weight of 1 and the other option received a weight of 0.

Relationship between reward ratios in each modality and the probability of choice

We found a weak main effect of modality ($F[2,38] = 5.95$, $p = 0.024$) on choice ratios, indicating that choice ratios differed between modalities. This effect corresponded to a tendency of participants to choose the visual option more often than the auditory option in the audio-visual block even when they had the same reward ratio 1:1 as can be seen in Figure 2A, thereby creating a difference between choice ratios of intra- and inter-modal conditions. However, this difference only reached significance for the reward ratio 3:1 as in inter-modal trials as participants chose the auditory modality significantly less often than options in intra-modal trials (Table S1). Note that since LNP models fitted to the fMRI data were estimated for each condition separately, this bias (i.e., preference of visual over auditory stimuli in audiovisual blocks) does not have any impact on our reported results regarding the differences of value representations between modalities.

Table S2. Results of the post-hoc pairwise comparisons of choice ratios between different modalities.

Reward Ratio	Modality-1	Modality-2	Difference	pValue
1:1	Auditory	Visual	0.0417	0.9857
1:1	Auditory	AudioVisual	0.1069	0.0626
1:1	Visual	Auditory	-0.0417	0.9857
1:1	Visual	AudioVisual	0.0651	0.3084
1:3	Auditory	Visual	0.0548	0.3278
1:3	Auditory	AudioVisual	0.0701	0.2903
1:3	Visual	Auditory	-0.0548	0.3278
1:3	Visual	AudioVisual	0.0154	1.0
3:1	Auditory	Visual	0.0325	1.0
3:1	Auditory	AudioVisual	0.1424	0.0873
3:1	Visual	Auditory	-0.0325	1.0
3:1	Visual	AudioVisual	0.1098	0.0156*

* Indicates significance at $p<0.05$

Analysis of the reaction times (RTs)

Similar to the analysis of the absolute differential value, we analysed the mean reaction time (RT) data for the two types of feedback in the value and the control tasks. A two-way repeated-measures ANOVA of RTs with task and feedback as factors revealed no significant main or interaction effects (p-values > 0.05). However, a trend was found for the main effect of task $F[2,38] = 4.76$, $p = 0.06$, reflecting faster responses in the control compared to the value task. Overall, the mean RT in the control task (787.5(±0.0613) ms), where participants had to simply follow instructions for decision-making was shorter than the mean RT in value task (824.4(±0.0635) ms).

Intuitively, a systematic decrease in the mean RTs of value task along with an increase in the absolute differential values (from no-reward to reward feedbacks), would indicate that participants take more time to reach a decision during difficult choice trials (when both options were perceived as having approximately equal values) in comparison to easy choice trials (when one option was clearly more valuable than the other). In the value task, mean (±s.e.m.) RTs decreased from 828.3(±0.0634) ms (no reward/blue feedbacks) to 819.7(±0.0638) ms (reward/yellow feedbacks). On the contrary, in the control task, mean RTs increased from 786.7(±0.0620) ms (keep/blue feedbacks) to 792.9(±0.0580) ms (switch/yellow feedbacks). Although insignificant, the latter effect implies an obvious fact that participants took less time when they had to keep their past choice in comparison to making a switch. Neither the main effect of feedback type on RTs nor their interaction with the task however reached significance, based on ANOVA ($Fs<1$, $p>0.1$).

Comparison of the value and the control task

In addition to the value task, we also inspected the control task using the same contrasts that were used to detect modality-specific and modality-general representations shown in **Figure 3**. Interestingly, we found that in the control task there were weak activations (**Figure S4** and **Table S2**) in vmPFC that overlapped with modality-general representations that were found in the value task. This observation indicates that a task with comparable choice structure but no valuation requirement also involves vmPFC, underscoring the role of this region as a general comparison and choice computation region.

Table S3. Univariate contrasts inspected in the control task

Contrast	Region	*x*	*y*	*z*	*t*(19)	*k*	*SVFWE corr P*
intraaudSV > 0	vmPFC	-8	62	-4	4.49	11	0.300
intravisSV > 0	vmPFC	-6	64	-6	4.37	12	0.287
interaudvisSV > 0	vmPFC	-4	56	-12	2.87	8	0.966

MNI coordinates (x, y, z) and T value corresponds to the local maxima peak of the cluster activations at SVFWE corrected P (cluster labels are from AAL atlas (Rolls et al., 2020)).

In order to rule out that the functional specialization of OFC clusters for visual and auditory value is due to the differences in the sensory properties of these stimuli, we explicitly compared the two tasks against each other by measuring the interaction between value modulations in each modality and task. Specifically, the differential contrasts of all auditory domain regressors of the value task against the control task (*ValueAud* > *ControlAud*) and all visual domain regressors of the value task against the control task (*ValueVis* > *ControlVis*) were inspected. This analysis revealed the modality-specific valuation clusters in left lateral and posterior OFC indicating that these clusters had significantly higher activations in the value compared to the control task (**Table S4**). However, the interaction contrasts revealed no activation in modality-general regions identified in vmPFC, again indicating that vmPFC plays a general role in the final choices but not the processing of stimulus value.

Table S4. Modality-specific stimulus value clusters in OFC for various differential interaction contrasts

Contrast	Region	*X*	*Y*	*Z*	*t*(19)	*k*	*SVFWE corr*
ValueAud > *ControlAud*	latOFC_L	-52	24	-6	6.34	63	P < 0.005
ValueVis > *ControlVis*	postOFC_L	-36	24	-18	4.91	32	P = 0.096

MNI coordinates (x, y, z) and T value corresponds to the local maxima peak of the cluster activations at SVFWE corrected (cluster labels are from AAL atlas (Rolls et al., 2020)).

Effective Connectivity Analysis

The model space shown in Figure 4 and 5 was developed over a base model comprising driving inputs and intrinsic connections, which did not vary with the experimental conditions. The rest of the models differed from each other over modulatory connections, which depended on the experimental conditions. In the base model, intrinsic connections were defined between every pair of nodes in the network and as self-connections. In Table S4 and S5, we show the estimated strength of the intrinsic connectivity (Table S4) and driving inputs (Table S5) of the winning model.

Table S5. Intrinsic connectivity parameters in the network including self-connectivity

To\From	vmPFC	audOFC	visOFC	audSen	visSen
vmPFC	-1.4645	-0.0433	0.2394	-0.0680	-0.0310
audOFC	-0.1555	-1.1751	0.2242	-0.0215	0.0809
visOFC	0.0250	-0.0477	-0.8745	0.0015	0.0514
audSen	0.0491	-0.0508	-0.0251	-0.9244	0.0671
visSen	-0.0213	0.0334	-0.0159	0.0005	-1.1425

All parameters are significant at posterior probability of P > 0.99

Table S6. Driving input influence parameters on sensory ROIs of the network

ROIs\Driving Input*	intraAud	intraVis	interAud	interVis
audSen	0.0042	0	0.0226	0.0224
visSen	0	-0.0025	-0.0035	0.0240

All parameters are significant at posterior probability of P > 0.99
*For trial-types of different conditions of the value task as described in the methods section on EC

Citations

Rolls ET, Huang CC, Lin CP, Feng J, Joliot M (2020) Automated anatomical labelling atlas 3. Neuroimage 206:116189.

Rolls ET, Joliot M, Tzourio-Mazoyer N (2015) Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. Neuroimage 122:1–5.