

Appetitive and Aversive Goal Values Are Encoded in the Medial Orbitofrontal Cortex at the Time of Decision Making

Hilke Plassmann,^{1,2} John P. O'Doherty,^{3,4,5} and Antonio Rangel^{3,4}

¹Institut Européen d'Administration des Affaires, 77305 Fontainebleau, France, ²Inserm, Ecole Normale Supérieure, 75005 Paris, France, ³Humanities and Social Sciences and ⁴Computational and Neural Systems, California Institute of Technology, Pasadena, California 91125, and ⁵Institute of Neuroscience, Trinity College, Dublin 2, Ireland

An essential feature of choice is the assignment of goal values (GVs) to the different options under consideration at the time of decision making. This computation is done when choosing among appetitive and aversive items. Several groups have studied the location of GV computations for appetitive stimuli, but the problem of valuation in aversive contexts at the time of decision making has been ignored. Thus, although dissociations between appetitive and aversive components of value signals have been shown in other domains such as anticipatory and outcome values, it is not known whether appetitive and aversive GVs are computed in similar brain regions or in separate ones. We investigated this question using two different functional magnetic resonance imaging studies while human subjects placed real bids in an economic auction for the right to eat/avoid eating liked/disliked foods. We found that activity in a common area of the medial orbitofrontal cortex and the dorsolateral prefrontal cortex correlated with both appetitive and aversive GVs. These findings suggest that these regions might form part of a common network.

Introduction

A sizable fraction of human and animal decisions involves value-based choice (Balleine and Dickinson, 1998; Daw et al., 2005; Rangel et al., 2008). In this type of choice, individuals begin the decision process by first assigning values to all of the stimuli that can be obtained, and then comparing the computed values to select one of them. The values assigned to the different stimuli at the time of decision making are often called “goal values.”

An important open question is whether different types of choice situations engage different valuation networks. For example, individuals sometimes have to choose between multiple appetitive items, as epitomized by the decision of a hungry diner facing a well stocked buffet table. But at other times they need to choose between multiple aversive items, for example, when choosing which of two undesirable risks to take. In this paper, we investigate whether goal values in these two cases are computed in the same regions or whether, in contrast, distinct neural substrates are at work.

There are considerable behavioral and neurobiological reasons to take seriously the hypothesis that there might be two goal valuation systems, one for appetitive decisions and one for aversive ones. First, evidence for the existence of neuroanatomically distinct appetitive and aversive responses has been found in the representation of other decision-related signals such as “outcome values,” “anticipatory values,” and “prediction errors” (O'Doherty et al., 2001; Rolls et al.,

2003; Small et al., 2003; Seymour et al., 2005; Ursu and Carter, 2005; Yacubian et al., 2006; Liu et al., 2007; Elliott et al., 2010). Second, a large body of evidence in behavioral economics suggests that valuation functions in the appetitive domain (gains) have different properties from those in the aversive domain (losses) (Kahneman and Tversky, 1979). In particular, in the case of monetary choices, it has been shown that individuals weight losses more strongly than gains of a similar magnitude. A natural explanation for such an asymmetry is the existence of a separate valuation system for gains and losses.

There are also reasons to consider the possibility that appetitive and aversive goal values might be encoded in overlapping areas. Although several monkey electrophysiology (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006) and human functional magnetic resonance imaging (fMRI) studies (Arana et al., 2003; Paulus and Frank, 2003; Plassmann et al., 2007; Chib et al., 2009; Hare et al., 2008, 2010) have found activity in medial orbitofrontal cortex (mOFC) associated with appetitive goal values, Tom et al. (2007) found a decrease in activity in medial OFC while subject evaluated gambles involving monetary loss. However, because in that study the “aversive” outcomes were given by a reduction in the overall monetary payoff from participating in the experiment, the prospect of monetary loss might not have activated the aversive system. Thus, the means by which aversive goals are evaluated in the brain at the time of decision making has not been fully addressed. In this paper, we report the results of two fMRI studies in which subjects made choices to avoid eating aversive food outcomes to provide additional insight into this question.

Materials and Methods

The first study looked for areas that might encode aversive goal values. In particular, we scanned satiated subjects' brains using fMRI while they repeatedly placed real bids for the right to avoid eating different disliked

Received Feb. 12, 2010; revised June 15, 2010; accepted June 27, 2010.

This work was supported by National Science Foundation Grant SES-0134618 and by a grant from the Gordon and Betty Moore Foundation to the Caltech Brain Imaging Center.

Correspondence should be addressed to Antonio Rangel, California Institute of Technology, MC 228-77, Pasadena, CA 91125. E-mail: rangel@hss.caltech.edu.

DOI:10.1523/JNEUROSCI.0788-10.2010

Copyright © 2010 the authors 0270-6474/10/3010799-10\$15.00/0

foods (e.g., baby foods, canned meat) in a Becker–DeGroot–Marshak (BDM) auction (Becker et al., 1964) (Fig. 1). A key feature of the design was that the bids entered by subjects generated reliable real-time measures of the aversive goal values assigned to every item: the higher the bid, the more aversive the item. We could then use the bids to look for brain areas that parametrically correlated with aversive goal values. Another important feature of the design was the presence of two kinds of trials: free-bid trials and forced-bid trials. These two trial types were identical except that, whereas subjects were free to select the amount of their bid in the free trials, they were required to bid a certain amount in the forced trials. As explained in more detail below, this allowed us to separate areas that encode for goal values from areas that encode for other correlated but different computations (e.g., anticipatory disgust).

The second study investigated whether the same overlapping areas of mOFC and dorsolateral prefrontal cortex (DLPFC) are involved in appetitive and aversive goal value computations. We performed a follow-up study during which subjects made choices involving appetitive and aversive items on separate trials. In particular, subjects now could enter positive bids to indicate that the items were appetitive, and if so how much, and negative bids to indicate that the items were aversive, and if so how much. All other details of the experiment remained unchanged. In addition to providing an independent replication test of the first study and of Plassmann et al. (2007), a key feature of this second study was that it allowed us to compare directly between neural representations of aversive and appetitive goal values within the same sample of subjects, rather than just performing such comparisons in a qualitative manner between different subject groups and across studies. Note that goal values refer to the value assigned to stimuli at the time of choice, in some choice paradigms involving learning, they are also called “Q values.”

Study 1

Subjects. Nineteen normal-weight subjects participated in the experiment (15 males; mean age, 23.7; age range, 18–47). One additional subject participated in the experiment but was excluded from the analysis because of technical problems during data acquisition. All subjects were right-handed, were healthy, had normal or corrected-to-normal vision, had no history of psychiatric diagnoses, neurological or metabolic illnesses, and were not taking medications that interfere with the performance of fMRI. All subjects had no history of eating disorders and were screened for disliking the types of foods used in the experiment (subjects who had allergies or other medical reasons not to consume some of the foods were excluded from the experiment). Also, subjects who participated in the study by Plassmann et al. (2007) were not eligible to participate in this study.

Stimuli. Subjects bid on 50 different sweet and salty foods that, during pretesting, were rated as either neutral or aversive by our subject population. Examples of the foods used include various types of baby foods and canned meats. We selected the foods based on pilot data to satisfy several characteristics. First, we wanted items that were available at local supermarkets. Second, we wanted items to be aversive for our subjects (in the sense that they were willing to pay to avoid having to eat the items at the end of the experiment). The behavioral data show that subjects indeed disliked the foods: the average willingness-to-pay (WTP) to avoid having to eat the items was \$1.75 (SE, 0.17), and the bids were negatively correlated with ratings of how much subjects would enjoy eating the foods collected at the end of the experiment ($r = -0.77$; $p < 0.0001$) (see Fig. 2c).

The foods were presented to the subjects using high-resolution color pictures. The stimulus presentation and response recording was con-

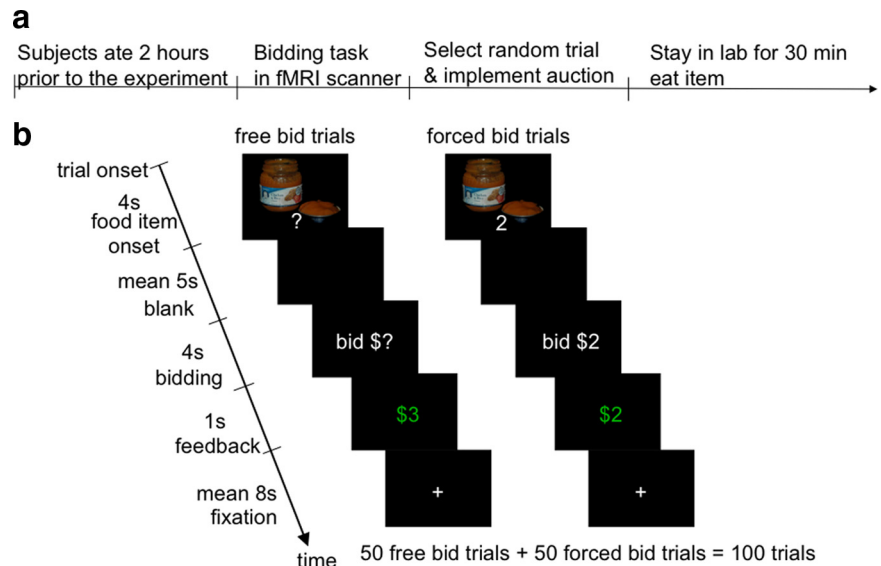


Figure 1. Study 1: experimental design. *a*, Timeline of the experiment. *b*, Time course for free- and forced-bid trials. Free- and forced-bid trials were identical except that, in forced-bid trials, subjects choose how much to bid (\$0, \$1, \$2, or \$3) for the right not to have to eat the food item, and in forced-bid trials they were forced to bid the amount shown with the food stimulus. Food items, trial type, and forced-bid amounts were fully randomized within subjects. The design for study 2 is very similar except that subjects are shown both appetitive and aversive items and choose a bid in the range -3 to 3 to indicate their willingness to pay to consume the item (which is negative for aversive items).

trolled by E-prime (Psychology Software Tools). The visual stimuli were presented using video goggles.

Task. Figure 1 describes the time structure of the experiment. Subjects were instructed to eat at most 2 h before the experiment to decrease the attractiveness of the foods. During the informed consent period, subjects agreed to eat whatever food item they would eventually have to eat as determined by the auction mechanism described below. Note that, although none of the subjects failed to comply with this instruction, they had the right to withdraw from the study at any point, in accordance with standard human subject protections.

Subjects were instructed that they would have to remain in the laboratory for 30 min at the conclusion of the experiment and that they will have to eat whatever food was shown on a randomly selected trial unless they purchased from us the right not to do so. In addition to a \$50/h participation fee, each subject received three \$1 bills in “spending money” to place bids to purchase the right to avoid eating disliked food items. Whatever money they did not spend was theirs to keep.

Subjects placed bids for the right to avoid eating particular foods in 100 different bidding trials. In each trial, they were allowed to bid \$0, \$1, \$2, or \$3. At the end of the experiment, one of those trials was randomly selected, by drawing a ball from an urn, and only the outcome of that trial was implemented. As a result, subjects did not have to worry about spreading their \$3 budget over the different items and they could treat each trial as if it were the only decision that counted. Objects were sold using a modified version of the Becker–DeGroot–Marshak auction. The rules of the auction were as follows. Let b denote the bid made by the subject for a particular item. After the bid is made, a random number n was drawn from a known distribution (in our case, \$0, \$1, \$2, and \$3 were chosen with equal probability). If $b \geq n$, the subject did not have to eat the item and paid a price equal to n . In contrast, if $b < n$, the subject had to eat the entire portion of the food shown but also did not have to pay anything. The only reason allowed for not eating the food if the auction was lost were medical allergies. No subject refused to eat the food in the experiment. Subjects were aware of these instructions from the beginning of the experiment (instructions are included in the supplemental material, available at www.jneurosci.org).

We used this auction because it has three very useful properties. First, it is characterized by a simple set of rules. Second, the optimal strategy for a buyer is to bid exactly her WTP for avoiding the food item being sold.

The intuition for why this is the case is as follows. There is no incentive to bid less than the WTP since the price paid is determined by the random number n , and thus the bids do not affect the price paid. There is also no incentive to increase the bid above the WTP since this may lead to a situation in which the subject avoids having to eat the item but ends up paying a price larger than his WTP (e.g., consider the case of WTP = \$1, $b = \$3$, and $n = \$2$). The fact that bidding the WTP is the optimal strategy was explained and emphasized extensively during the instruction and training period. We performed extensive pilot work to find a set of instructions that lead to 100% reported compliance with the best strategy. The instructions emphasized that the subject's best strategy is to look at the item, ask herself how much avoiding to eat the food was worth to her, and simply bid that amount. Third, and most importantly, since bidding the exact WTP is the best strategy, the bids provide a good measure of the aversive goal value computed by the brain for every subject and item at the time of decision making. Thus, we could use the bids as a behavioral measure of aversive goal value to correlate with the blood oxygen level-dependent (BOLD) signal on a trial-by-trial basis.

It is important to highlight several limitations of this procedure for measuring the values. First, to keep the task simple in the scanner, subjects were only allowed to bid discrete amounts (\$0, \$1, \$2, or \$3). A consequence of this is that the bids were only approximations of the true WTP computed by subjects. For example, when the true WTP is \$2.3, our measure is \$2. Similarly, subjects with a WTP larger than \$3 enter a bid of \$3. However, the data suggest that the approximation is a good one (see Fig. 2c). Second, theoretical work (Karni and Safra, 1987; Horowitz, 2006) has shown that bidding the WTP is a dominant strategy only when individuals have expected-utility like preferences, and that otherwise the optimal bid depends on the underlying distribution prices used by the BDM procedure, an effect that has been shown in experimental work (Bohm et al., 1997). This raises the possibility that there might have been an additional source of noise in our measurement of the values. However, the theoretical work shows that the bids preserve the monotonicity of the underlying values even under unexpected utility preferences.

We used two different kinds of trials: free-bid trials and forced-bid trials. Each of the 50 items was shown twice, once in a free trial and once in a forced trial. The order of these trials was fully randomized across subjects and food items. Both types of trials had equal probability of being selected to be the trial that was implemented at the end of the experiment. The timing for each type of trial is shown in Figure 1b. Subjects were required to enter a bid within the 4 s bidding window. Trials in which subjects failed to do so were classified as missing trials in the analysis described below.

A key difference between the two types of trials is that, whereas subjects were free to select the amount of their bid in the free trials, they were told how much to bid in the forced trials. The forced bids were drawn uniformly and independently from \$0, \$1, \$2, \$3 on each trial. The set of rules described above applied to both trials. As a result, subjects need to make an aversive goal value computation in free trials to decide how much to bid, but they do not necessarily need to do so in forced trials.

The existence of both types of bidding trials is an essential component of the experimental design. A difficulty in searching for the neural basis of the aversive goal value computation is that, when the brain is exposed to a picture of an aversive food item, it might simultaneously compute other variables that are correlated with goal value. For example, the brain may simulate the anticipated taste of the food, which may lead to a calculation of the level of disgust. If this issue is not properly addressed, one could erroneously attribute goal value computations to areas that are calculating different, albeit correlated, variables. The presence of free and forced trials provides a potential solution to the problem: the subject needs to perform a goal value computation in the free trials, since she needs to decide how much to bid, but not in the forced trials, since she is told what her bid should be. As a result, any area that encodes for goal values in free, but not in forced trials, is a good candidate for the goal value computation. In contrast, areas whose activity correlates with goal values in both types of trials might be encoding goal values or different, but correlated, computations. Fortunately, all of the key areas of interest identified in both studies have the property that they are correlated with goal values in

free but not in forced trials, so this potential false negative is not an issue. One caveat to this procedure is that subjects might engage in more reliable anticipatory computations of goal value in free than in forced trials, since they know that they have more control over their bids. This fact, together with the imposition of randomly selected bids in the forced trials, could partially contribute to the difference in correlation with the value signals in free and forced trials.

After receiving the instructions (included in the supplemental material, available at www.jneurosci.org), subjects were trained on the use of the response boxes with their right hand and on the bidding procedure. To avoid activation artifacts attributable to the assignment of buttons to bid amounts, the assignment was counterbalanced across subjects.

After the scanning session, subjects performed a liking-rating task. Subjects were asked to rate "how much would you like to eat this food at the end of the experiment" for each of the items. We used a scale from 1 (not at all) to 5 (very much). Each food picture was shown in random order and subjects had an unlimited amount of time to enter their rating.

fMRI data acquisition. The functional imaging was conducted using a Siemens 3.0 tesla Trio MRI scanner to acquire gradient echo T2*-weighted echo-planar (EPI) images with BOLD contrast. To optimize functional sensitivity in OFC, we used a tilted acquisition in an oblique orientation of 30° to the anterior commissure–posterior commissure line. In addition, we used an eight channel phased array coil, which yields a 40% signal increase in signal in the medial OFC over a standard head coil. Each volume comprised 32 axial slices. A total of 1100 volumes (two sessions; ~18 min each) were collected during the experiment in an interleaved-ascending manner. The imaging parameters were as follows: echo time, 30 ms; field of view, 192 mm; in-plane resolution and slice thickness, 3 mm; repetition time, 2 s; no interslice gap. Whole-brain high-resolution T1-weighted structural scans (1 × 1 × 1 mm) were acquired from the 19 subjects and coregistered with their mean EPI images and averaged together to permit anatomical localization of the functional activations at the group level. Image analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Temporal normalization was applied to the scans with an acquisition time of 1.9375 referenced to the last volume. To correct for subject motion, the images were realigned to the last volume, spatially normalized to a standard T2* template with a resampled voxel size of 3 mm, and spatially smoothed using a Gaussian kernel with a full width at half-maximum of 8 mm. Intensity normalization and high-pass temporal filtering (using a filter width of 128 s) were also applied to the data.

fMRI data analysis. Two different, albeit closely related, general linear models were used in the paper. The two models are described in detail in the supplemental material (available at www.jneurosci.org). Here, we briefly sketch the characteristics of the main model used to generate the results described in the paper.

The data analysis proceeded in three steps. First, we estimated a general linear model with AR(1) and the following regressors that capture the main events in our experiment: R1, free bid at picture presentation; R2, free bid at response; R3, forced bid at picture presentation; R4, forced bid at response; R5, missed bid trial and picture presentation; and R6, missed bid trial at response. The regressors that capture the presentation of the food pictures were modeled using 4 s events. The regressors for the bid responses were modeled as events with a duration equal to the response time of the subject (measured from the appearance of the bid screen). The regressors for the missed bid trials were modeled as 4 s events. To take advantage of the parametric nature of our design, we included two modulators: the first regressor was modulated by the magnitude of the free bid and the third regressor was modulated by the bid that the item shown in that trial received in the free condition. These modulators were constructed subject-by-subject. Note that the free bid is a measure of the subject's aversive goal value for the item being shown. We also included a constant term and six motion parameters as regressors of no interest for each session. Each of the regressors was convolved with a canonical hemodynamic response function.

Second, we calculated the following first-level single-subject contrasts: (1) free-bid trials while exposed to item modulated by free bid (regressor R2), (2) forced-bid trials while exposed to item modulated by its respective free bid (regressor R4), and (3) regressors R2 minus R4.

Third, we computed group random-effect estimates using the whole-brain and region-of-interest (ROI) techniques. For the whole-brain analyses, we calculated random-effects group estimates contrasts using one-sample *t* tests over the single-subject contrast statistics. For inference purposes, we used a threshold of $p < 0.001$, uncorrected, and an extent threshold of 5 voxels. For expositional purposes only, all of the activation maps in the paper are drawn at the more liberal threshold of $p < 0.005$, uncorrected, and an extent threshold of 5 voxels. The ROI analyses were based on a priori hypotheses about five areas that have been previously shown to be involved in different aspects of valuation in related contexts: bilateral medial orbitofrontal cortex, right dorsolateral prefrontal cortex, bilateral insula, lateral orbitofrontal cortex, and ventral striatum. The ROI analysis was performed by carrying out one-sample *t* tests over the single-subject contrast statistics using an statistical threshold that was small volume and false discovery rate corrected at the $p < 0.01$ level. This correction was implemented with the SPM software. (For additional details, see supplemental material, available at www.jneurosci.org.)

Study 2

Since the experiment is very similar to study 1, we only describe the differences between them.

Subjects. Twenty normal-weight subjects participated in the experiment (15 males; mean age, 23–25; age range, 19–34). Two additional subjects participated in the experiment but were excluded from the analysis because of technical problems during data acquisition. In addition to all of the exclusion criteria listed above, subjects were screened for liking and disliking the types of foods used in the experiment. None of the subjects who participated in this experiment was part of study 1.

Stimuli. Subjects bid on 60 different sweet and salty foods. During pretesting, 30 of those items were rated as aversive and 30 as appetitive or neutral by our subject population. Examples of the aversive foods used include various types of baby foods and canned meats. Examples of the appetitive foods used include various types of candy bars and chips. We selected the foods based on pilot data to satisfy several characteristics. First, we wanted items that were available at local supermarkets. Second, we wanted items to span a close-to-uniform distribution of appetitive and aversive ratings.

Task. Subjects received the instructions to come neither full nor hungry to the experiment. As before, subjects were instructed that they would have to remain in the laboratory for 30 min at the conclusion of the experiment and that the food that they will have to eat would depend on their bidding behavior on a randomly selected trial. Instead of placing bids for the right not to have to eat a food, subjects now placed bids for the right to eat or to avoid eating particular foods in 120 different bidding trials. In each trial, they were allowed to bid $-\$3$, $-\$2$, $-\$1$, $\$0$, $\$1$, $\$2$, or $\$3$. Foods were sold using a modified version of the Becker–DeGroot–Marschak auction as described above. Note that positive bids (i.e., a willingness to pay) now indicated the value placed on consuming the item and that negative bids indicated the amount of money that the subject needs to be compensated to be willing to consume the item. Also, the random number used to determine the outcome of the auction now was randomly drawn from -3 to $+3$. The possibility of placing positive and negative bids does not change the key property of the BDM auction procedure: subjects' best strategy still is to bid the amount (in dollars) that they like or dislike the item. This implies that the bids were a good approximation to the appetitive or aversive goal value computed by the brain in every trial.

Each of the 60 items was shown twice, once in a free-bid trial and once in a forced bid trial.

fMRI data acquisition. Data acquisition of study 2 was identical with that of study 1 except that a total of 1473 volumes (three sessions; ~ 12 min each) was collected in an interleaved-ascending manner.

fMRI data analysis. The analysis proceeded exactly as in study 1. However, note one key difference between the two studies. In experiment 1, subjects bid for the right not to have to eat the item. As a result, higher bids mean more aversive goal values. In contrast, in experiment 2, subjects bid for the right to have to eat the item (if positive) or for the amount of compensation that they require to be willing to eat the item (if negative). In this case, larger bids mean more appetitive goal values.

Results

Study 1

Behavioral results

Figure 2*a* shows the distribution of bids during free- and forced-bid trials. The average bid in free trials was $\$1.75$ (SE, 0.17), which was not significantly different from the average bid on the forced trials, and $>75\%$ of the free bids were greater than zero. The bid amounts for the forced-bid trials were randomly drawn from a uniform distribution on $\$0$, $\$1$, $\$2$, and $\$3$.

Figure 2*b* shows the distribution of bid reaction times for free- and forced-bid trials by bid amount. Two-sided paired *t* tests showed significant differences between the free versus forced trials for $\$1$ bids, but not for the other amounts. The same test also showed a significant difference in reaction times between $\$1$ and $\$2$ free-bid trials and $\$0$ and $\$3$ free-bid trials, perhaps reflecting a difference in difficulty between extreme and nonextreme aversive goal values.

To verify that the bids in the free-bid trials were good measures of the aversive goal values, we also sampled liking ratings for each food after the scanning session (scale: 1, not at all, to 5, neutral). Figure 2*c* shows that the free bids correlated negatively with subjects' reported liking for the foods ($r = -0.77$; $p < 0.0001$). These findings suggest that the subjects' bids reflect their disliking for the foods and that subjects understood the bidding procedure described in Materials and Methods.

Imaging results

We estimated a general linear model to test for brain regions in which BOLD activity was correlated (positively or negatively) with the aversive goal values, as measured by the individual subject bids, during the evaluation period at the beginning of each trial (for details, see Materials and Methods and supplemental material, available at www.jneurosci.org). This model was used to identify three key contrasts of interest: (1) areas correlated with aversive goal values during the free-bid trials, (2) areas correlated with aversive goal values during the forced-bid trials, and (3) areas that exhibited a stronger correlation with goal values during the free-bid trials than during the forced-bid trials. As discussed in Materials and Methods, the last contrast is essential because it allows us to identify areas that are related to the computation of goal values, as opposed to areas that are involved in computations that are correlated with goal values but distinct from them (e.g., anticipatory disgust).

We performed two kinds of random-effects group analyses: whole-brain analysis (at a statistical threshold of $p < 0.001$, uncorrected, and an extent threshold of 5 voxels) and ROI analyses. These later analyses were based on a priori hypotheses about four areas that have been previously shown to be involved in different aspects of valuation in related contexts: the bilateral mOFC, the right dorsolateral prefrontal cortex (rDLPFC), the right lateral OFC (lOFC), the ventral striatum (VtStr), and the bilateral insula. These areas were selected as candidate regions for an aversive goal valuation system based on findings from the previous literature in behavioral neuroscience. All ROI statistics were small-volume false-discovery rate corrected (SVFDR) at the $p < 0.01$ level. In the results below (as well as in the supplemental tables, available at www.jneurosci.org as supplemental material), we report the results of both types of analyses.

We selected these five ROIs for the following reasons. First, we included the mOFC and the rDLPFC in our set of ROIs because a previous study by the authors, which used a similar design in an appetitive context (Plassmann et al., 2007), found that BOLD activity in these two areas was correlated with appetitive goal

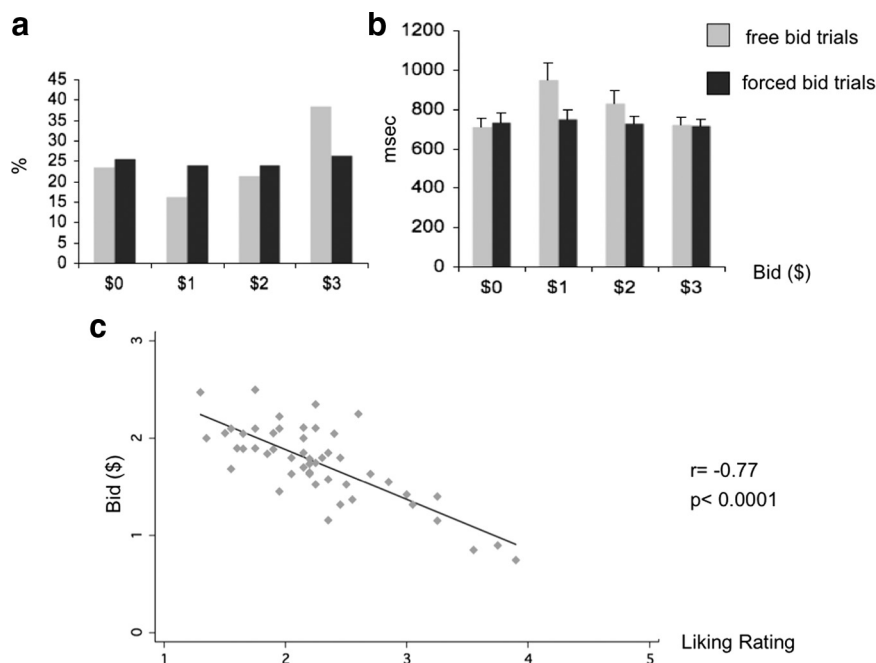


Figure 2. Study 1: behavioral results. *a*, Distribution of bids in free and forced trials. *b*, Reaction times for free- and forced-bid trials as a function of the size of the bid in that trial (bars denote SEs). *c*, Correlation and linear regression fit between liking ratings (sampled after the scanning session) and free bids (sampled during scanning session). Each point denotes a food item averaged across subjects.

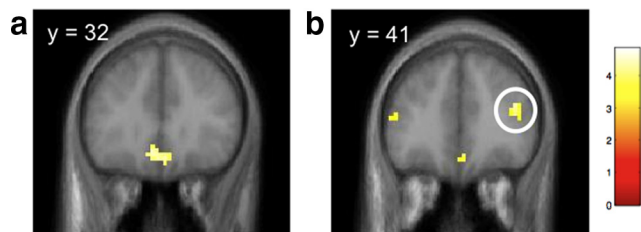


Figure 3. Study 1: random-effects ROI analyses shows that activity in the bilateral medial OFC ($x = -2, y = 32, z = -14, t = 3.79; x = 3, y = 32, z = -17, t = 3.41$) (*a*) and the bilateral DLPFC ($x = -46, y = 35, z = 14, t = 3.86; x = 39, y = 41, z = 22, t = 3.66$) (*b*) are negatively correlated with the bid in the free trials at the time of evaluation. The results are overlaid on the mean anatomical image. The color scale represents the magnitude of t values. For visualization purposes, we use a threshold of $p < 0.005$, uncorrected, and a 5 voxel extent threshold. SVFDR statistics are reported in the text.

values. Second, we included the right IOFC in our set of ROIs because previous studies of valuation at the time of outcome have found that BOLD activity in this area correlates with different types of aversive outcomes in certain contexts (O'Doherty et al., 2001; Anderson et al., 2003; Gottfried et al., 2003; Rolls et al., 2003; Elliott et al., 2008). Third, we included the ventral striatum because several studies have found activity in this area that correlated with behavioral measures of goal values (Kable and Glimcher, 2007; Knutson et al., 2007; Tom et al., 2007). However, note that previous work (Hare et al., 2008) has shown that BOLD activity in this area is involved in the encoding of prediction errors, but not of goal values at the time of decision making. Finally, we included the insula because previous studies have shown that BOLD activity in this area correlates with several aspects of disgust (Sanfey et al., 2003; Wicker et al., 2003), the experience of negative emotions (e.g., fear), and converging evidence implicates this region in appetitive reactions to food (Balentine and Dickinson, 1998; Jabbi et al., 2007).

Correlation with aversive goal values in free trials. Supplemental Table 1 (available at www.jneurosci.org as supplemental material) lists the results of this contrast for both types of analyses. Interestingly, no brain areas showed positive correlation with aversive goal values at a threshold of $p < 0.001$, uncorrected, with a 5 voxel extent threshold. However, the whole-brain analyses identified several regions in which the BOLD signal and the aversive values were negatively correlated: the middle and superior frontal gyrus, the anterior cingulate cortex, the superior temporal gyrus, the posterior insula, the mOFC, and the DLPFC. Not surprisingly, given that the ROI analyses amounts to using a lower statistical threshold on the same contrasts, this analysis identified correlations in the bilateral mOFC (Fig. 3*a*) ($x = -2, y = 32, z = -14, t = 3.79, p_{\text{SVFDR}} < 0.01; x = 3, y = 32, z = -17, t = 3.41, p_{\text{SVFDR}} < 0.01$), the right DLPFC (Fig. 3*b*) ($x = 39, y = 41, z = 22, t = 3.66, p_{\text{SVFDR}} < 0.01$), and the bilateral posterior insula ($x = -36, y = -7, z = 7, t = 3.38, p_{\text{SVFDR}} < 0.01; x = 42, y = -10, z = 4, t = 3.73, p_{\text{SVFDR}} < 0.01$).

In contrast, activity in the IOFC and VtStr was not significantly correlated with the size of the bids at this same SVFDR threshold.

Correlation with aversive goal values in forced-bid trials. As shown in supplemental Table 2 (available at www.jneurosci.org as supplemental material), the whole-brain analysis identified only two regions with a (positive) significant correlation: the brainstem and the middle temporal gyrus. Also, and in sharp contrast to the free trials, the ROI analysis did not find either a positive or negative significant correlation with aversive goal values during the forced-bid trials at a level of $p_{\text{SVFDR}} < 0.01$.

Correlation with aversive goal values in free versus forced trials. A limitation of the first contrast is that it identifies areas whose activity is correlated with aversive goal values, but also identifies areas that encode for variables that might be correlated with it, such as anticipatory disgust, but that are not involved in the computation of goal values. To address this potential confound, we took advantage of our experimental design and tested for regions that correlate with bids in free trials more than in forced trials (see supplemental Table 3, available at www.jneurosci.org as supplemental material). No brain areas showed more positive correlation in free than forced trials with aversive goal values at a threshold of $p < 0.001$, uncorrected, and a 5 voxel extent threshold. The whole-brain analyses identified several regions in which the BOLD signal exhibited more negative correlation with the bids in the free trials: the posterior cingulate cortex, the medial frontal gyrus, the superior temporal gyrus, the midbrain, the fusiform and parahippocampal gyri, the posterior insula, the mOFC, and the DLPFC. The ROI analysis showed that mOFC (Fig. 4*a*) ($x = 2, y = 32, z = -23, t = 3.12, p_{\text{SVFDR}} < 0.01$), rDLPFC (Fig. 4*c*) ($x = 45, y = 41, z = 25, t = 3.30, p_{\text{SVFDR}} < 0.01$), and the bilateral posterior insula ($x = -39, y = -1, z = -8, t = 3.36, p_{\text{SVFDR}} < 0.01; x = 42, y = -13, z = 1, t = 2.97, p_{\text{SVFDR}} < 0.01$) exhibited the desired correlation. However, the contrast was not significant in the IOFC or VtStr at a level of $p_{\text{SVFDR}} < 0.01$.

Since the mOFC and the DLPFC have been shown to correlate with appetitive goal values (Plassmann et al., 2007), we decided to explore further the computations made in these two areas using a *post hoc* analysis. We extracted measures of the average response in these two ROIs to high (\$2 and \$3) and low (\$0 and \$1) free-bid items during the evaluation period (for details, see supplemental material, available at www.jneurosci.org). Figure 4, *c* and *d*, shows the resulting parameter estimates for both areas in free-bid trials. In both cases, the activity in the free-bid trials decreased with the magnitude of the free bid.

Study 2

Behavioral results

Figure 5*a* shows the distribution of bids during free- and forced-bid trials. The average bid in free trials was $-\$0.21$ (SE, 0.12), which was not significantly different from the mean bid in the forced trials, and $>75\%$ of the free bids were smaller or greater than zero. The bid amounts for the forced-bid trials were randomly drawn from a uniform distribution on $-\$3$, $-\$2$, $-\$1$, $\$0$, $\$1$, $\$2$, and $\$3$. Two-sided paired *t* tests showed no significant differences in reaction times between positive versus negative free bids (mean_{pos}, 909.55 ms; SE_{pos}, 54.52 ms; mean_{neg}, 905.81 ms; SE_{neg}, 60.19 ms; $T_{(19)} = 0.19$; $p = 0.85$).

To verify that the bids in the free-bid trials were good measures of the aversive and appetitive goal values, we also sampled liking ratings for each food after the scanning session (scale: 1, not at all, to 5, very much). Figure 5*b* shows that the free bids correlated with subjects' reported liking for the foods ($r = 0.930$; $p < 0.0001$). These findings suggest that the subjects' bids reflected their liking and disliking for the foods and that subjects understood the bidding procedure described in Materials and Methods.

Imaging results

The analysis proceeded as in study 1. However, note one key difference between the two studies. In experiment 1, subjects bid for the right not to have to eat the item. As a result, higher bids represented more aversive goal values. In contrast, in experiment 2, subjects bid for the right to eat the item (if positive), or for the amount of compensation that they required to be willing to eat the item (if negative). In this case, larger bids mean more appetitive goal values and lower bids mean more aversive goal values.

Correlation with goal values in free trials. Supplemental Table 4 (available at www.jneurosci.org as supplemental material) lists the results of this contrast. The whole-brain analysis did not reveal any areas exhibiting a negative correlation with the goal values at a threshold of $p < 0.005$, uncorrected, with a 5 voxel extent threshold. However, this analysis identified several regions in which the BOLD signal and the goal values were positively correlated: the lingual gyri, the anterior cingulate, the superior frontal gyrus, the mOFC, and the DLPFC. The ROI analyses showed that activity in the bilateral mOFC ($x = 4, y = 30, z = -13, t = 4.55, p_{\text{SVFDR}} < 0.01$; $x = -3, y = 31, z = -12, t = 4.55, p_{\text{SVFDR}} < 0.01$) and the right DLPFC ($x = 49, y = 30, z = 27, t = 2.88, p_{\text{SVFDR}} < 0.01$) was correlated with the magnitude of the goal

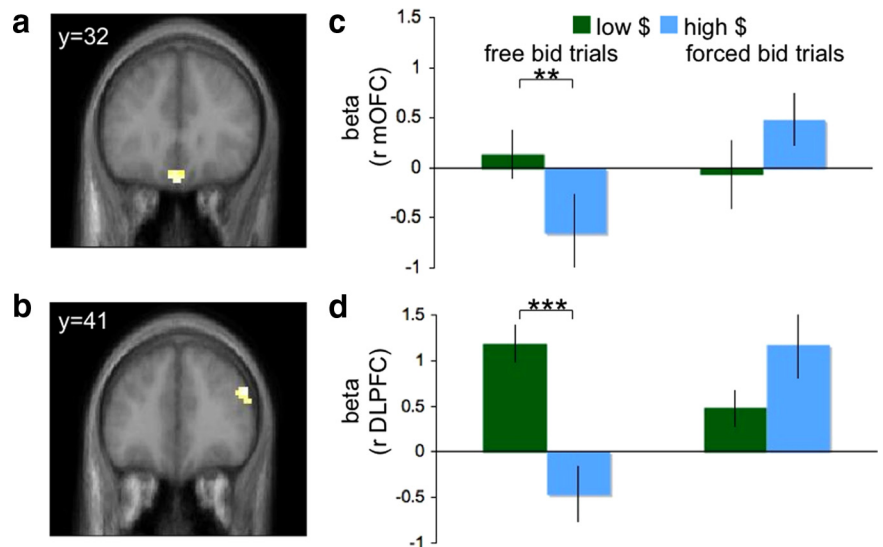


Figure 4. Study 1: random-effects ROI analyses shows that activity in the bilateral medial OFC ($x = 2, y = 32, z = -23, t = 3.12$) (*a*) and the right DLPFC ($x = 45, y = 41, z = 25, t = 3.30$) (*b*) are more negatively correlated with free bids during the free trials than during the forced trials at the time of evaluation. *c*, Estimates of the response in medial OFC to the stimulus screen for low free-bid ($\$0$ – $\$1$) versus high free-bid ($\2 – $\$3$) trials show that activity decreased with the bid in free-bid trials (paired *t* test, $p = 0.0027$) but not in forced-bid trials. *d*, Similar parameter estimates in right DLPFC show that activity decreased with the magnitude of the bid in free-bid trials ($p = 0.00001$) but not in forced-bid trials. Error bars indicate SEM. $**p < 0.005$; $***p < 0.001$.

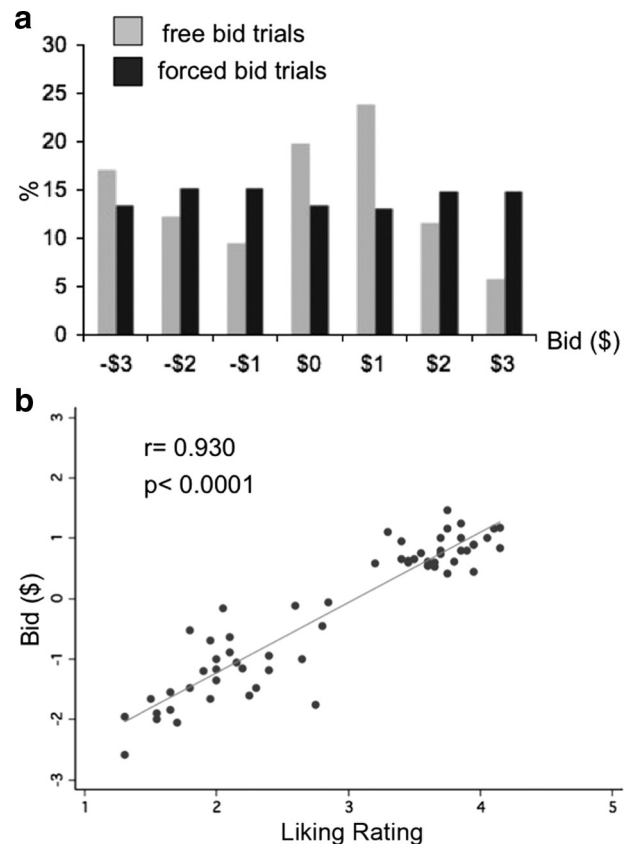


Figure 5. Study 2: behavioral results. *a*, Distribution of bids in free and forced trials. *b*, Correlation and linear regression fit between liking ratings (sampled after the scanning session) and free bids (sampled during scanning session). Each point denotes a food item averaged across subjects.

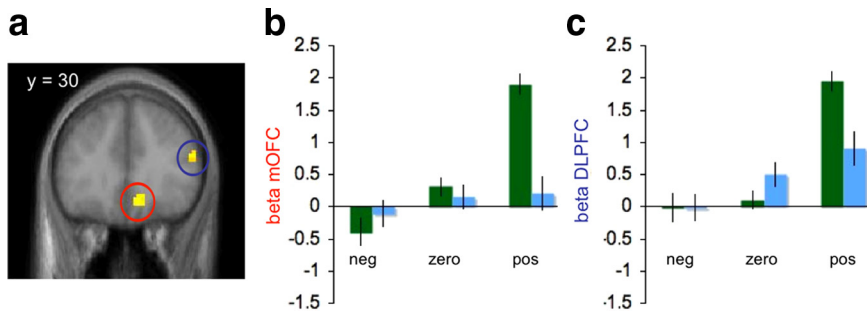


Figure 6. Study 2. *a*, Random-effects ROI analyses show that activity in the right medial OFC (red; $x = 6, y = 30, z = -12, t = 3.48$) and the right DLPFC (blue; $x = 50, y = 30, z = 21, t = 3.38$) are negatively correlated with the bid in the free trials at the time of evaluation. *b*, Estimates of the response in medial OFC to the stimulus screen for negative free-bid ($-\$3$ – $-\$1$), zero free bids, and positive free-bid ($\$1$ – $\$3$) trials in free-bid trials (green) show that activity increased with the bid, but not in forced-bid trials (blue). *c*, Similar parameter estimates in right DLPFC show that activity decreased with the magnitude of the bid. Error bars indicate SEM.

values. But at this same SVFDR corrected threshold there were no significant correlations in the insula, the IOFC, or the VtStr.

Correlation with goal values in forced-bid trials. Supplemental Table 5 (available at www.jneurosci.org as supplemental material) lists the results of this contrast. A whole-brain analysis showed that activity changes in the cuneus and lingual gyrus were positively correlated with goal values, and that activity in the parahippocampal gyrus and the superior temporal gyrus was negatively correlated with the same measure. Importantly, however, in none of our ROIs was activity significantly correlated with goal values at a level of $p_{SVFDR} < 0.01$.

Correlation with goal values in free versus forced trials. Supplemental Table 6 (available at www.jneurosci.org as supplemental material) describes the results of this contrast. A whole-brain analysis showed that activity in the parahippocampal gyrus, inferior temporal gyrus, mOFC, and DLPFC exhibited a larger positive correlation with goal values in free than in forced trials. The associated ROI analysis confirmed the activations in mOFC (Fig. 6*a*) ($x = 6, y = 30, z = -12, t = 3.48, p_{SVFDR} < 0.01$) and DLPFC (Fig. 6*a*) ($x = 50, y = 30, z = 21, t = 3.38, p_{SVFDR} < 0.01$) but found no significant excess correlations in IOFC, the insula, or the VtStr at the same significance level.

As before, since the mOFC and the DLPFC have been shown to correlate with appetitive goal values (Plassmann et al., 2007) and seemed also to be involved in aversive valuation in study 1, we decided to explore further the computations made in these two areas using a *post hoc* analysis. We extracted measures of the average response in these two ROIs during the evaluation period for positive (bid > 0), neutral (bid = 0), and negative (bid < 0) items. Figure 6, *b* and *c*, shows the resulting parameter estimates for each of the two areas for free- and forced-bid trials. In all three cases, the activity in the free-bid trials decreased with the magnitude of the free bid. Interestingly, activity in the mOFC and DLPFC during free trials was negative for aversive items and positive for positive items, which is consistent with the data from Plassmann et al. (2007) and study 1.

Correlation with goal values in free appetitive trials versus free aversive trials. The previous results show that, in study 2, activity in mOFC and DLPFC correlated with goal values in both the appetitive (non-negative bids) and aversive (negative bids) range. A natural question is whether there are areas that exhibit a stronger correlation with goal values in the positive or negative domains, which would point to some degree of specialization in appetitive or aversive goal value coding. To address this question, we estimated an additional general linear model (for

details, see supplemental material, available at www.jneurosci.org) in which we looked for areas that correlated with goal values more strongly during the free trials in either the non-negative or in the negative range. As reported in supplemental Table 7 (available at www.jneurosci.org as supplemental material), the bilateral insula, anterior cingulate cortex, and medial frontal gyrus exhibited stronger correlation with the bids during the non-negative trials. No areas exhibited stronger correlation with bids during the negative trials at our omnibus threshold.

In addition, we also look for areas that exhibited stronger excess correlation in free trials than in forced trials for either appetitive or aversive items. The results, described in supplemental Table 10 (available at www.jneurosci.org as supplemental material), show that this excess correlation is stronger in the insula for appetitive items, and in the inferior frontal gyrus for aversive ones.

Conjunction analysis

We concluded the analysis by performing a conjunction analysis (for details, see supplemental material, available at www.jneurosci.org) to identify regions in which brain activity was positively correlated with appetitive goal values in free trials in the study by Plassmann et al. (2007), negatively correlated with aversive goal values in free trials in study 1, and positively correlated with goal values (both aversive and appetitive) in free trials in study 2. The results, depicted in Figure 7, reveal overlapping loci of activation in the mOFC ($x = 3, y = 36, z = -18, t = 4.43, p_{SVFDR} < 0.01$) and the right DLPFC ($x = 42, y = 42, z = 15, t = 4.29, p_{SVFDR} < 0.01$).

Discussion

To make sound decisions, the brain needs to encode a number of distinct valuation signals (Rangel et al., 2008). First, at the time of evaluation, it needs to assign a value to the different options under consideration. These so-called goal value signals are used to guide choices. Second, it needs to encode the value of the outcome that it expects to receive following the selected option (“anticipatory values”). Third, the brain needs to measure the value of the actual outcome when it occurs (“outcome values”). The outcome value, together with the anticipatory value measure, is then used to compute a prediction error that measures the deviation between actual and anticipated values. Prediction errors serve as a teaching signal that can be used to improve future value assignments at the time of decision making.

Importantly, all of these value signals can be either positive or negative. For example, the outcome signals are positive for appetitive items that individuals like, and negative for aversive items that they dislike. This gives rise to a basic question in the neurobiology of decision making: Is the same region used to encode both valences of the different types of value signals, perhaps by encoding appetitive values with increases of activity, and aversive values with decreases of activity? Or, alternatively, are there dissociable areas encoding appetitive and aversive values? This question has received a considerable amount of attention in the case of outcome values, prediction errors, and anticipatory values, but not in the case of the computation of goal values at the time of decision making. One exception is a study by Tom et al. (2007)

that used fMRI to look for correlates of value signals for choices between monetary gambles. From the viewpoint of testing for the existence of separate goal value systems, a potential concern is that monetary payoffs might not have been able to engage a putative purely aversive valuation region in the same way that a biological punishment such as an unpleasant food stimulus would. In contrast to the study by Tom et al., in our studies subjects had to eat disliked foods (a real biological punisher) that provided us with a better chance of activating the hypothesized aversive valuation areas.

With respect to outcome valuation, correlates of positive outcome values have been found using fMRI in the medial OFC for odor (Anderson and Sobel, 2003; Rolls et al., 2003), flavor (Rolls et al., 2003; Small et al., 2003; McClure et al., 2004), attractive faces (O'Doherty et al., 2003b; Cloutier et al., 2008), and even secondary reinforcers such as money (Knutson et al., 2001; O'Doherty et al., 2001; Ursu and Carter, 2005; Liu et al., 2007; Elliott et al., 2010). Analogous results have been found for negative experiences: subjective reports of pain intensity correlate with activity in the insula and the anterior cingulate cortex (Peyron et al., 1999), and monetary losses, unattractiveness in face stimuli, and the presentation of aversive odors have been found to elicit activity in the lateral OFC (O'Doherty et al., 2001; Ursu and Carter, 2005; Liu et al., 2007; Elliott et al., 2010). Not all of the evidence, however, points in the direction of a dissociation between appetitive and aversive outcome value signals (Breiter et al., 2001; O'Doherty et al., 2003a; Dillon et al., 2008; Elliott et al., 2010). For instance, Elliott et al. (2003) found in a block design that activity in medial and lateral OFC was correlated with both wins and losses. O'Doherty et al. (2003) found that lateral OFC was engaged after detection of contingency changes after a monetary loss rather than by a monetary loss per se.

Another key valuation signal is the prediction error that encodes the difference between expected and actual outcomes, and that may play a role in the acquisition of value predictions. There is much evidence to implicate the phasic activity of dopamine neurons in encoding a prediction error signal during learning with reward stimuli (Schultz, 1998), and moreover human imaging studies report reward-prediction error-related activity in prominent target areas of these neurons in the striatum (Delgado et al., 2000; Pagnoni et al., 2002; Elliott et al., 2003; Yacubian et al., 2006; Seymour et al., 2007). The extent to which dopamine is also involved in responding to aversive events has been controversial (Horvitz, 2000). However, a recent neurophysiological investigation found evidence that some dopamine neurons responded to both appetitive and aversive stimuli (Matsumoto and Hikosaka, 2009), suggesting that these neurons may potentially contribute to the learning of aversive as well as appetitive associations.

Our results contribute to this literature by showing that a complete appetitive–aversive spatial dissociation does not seem to exist in the case of the goal valuation signal at the time of decision making. Instead, we found that activity in the mOFC

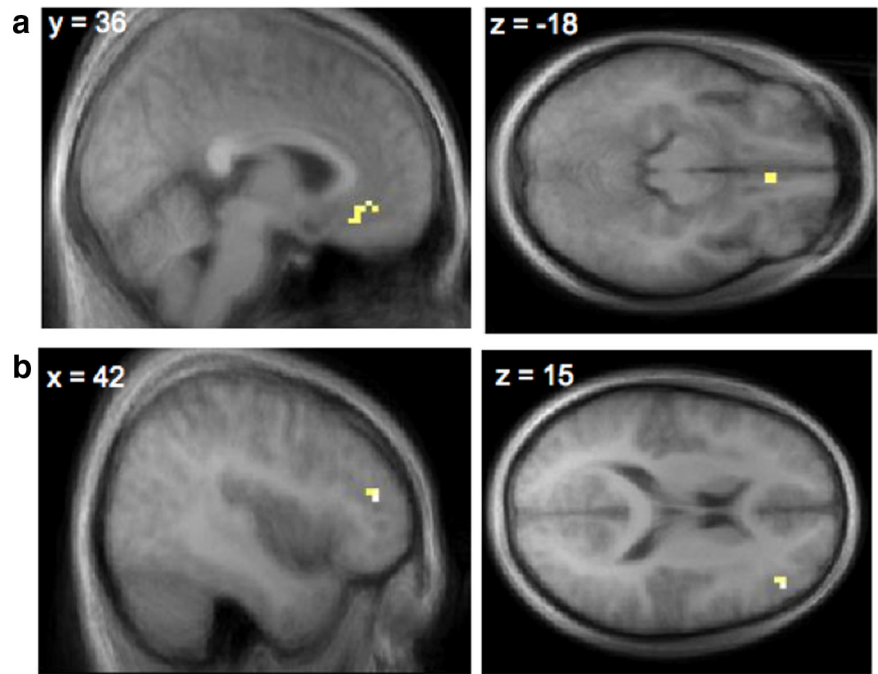


Figure 7. Comparison across studies. Results of a conjunction analysis to determine overlaps between areas that correlate positively with appetitive goal values during free-bid trials in the study by Plassmann et al. (2007) correlate negatively with aversive goal values during free-bid trials in study 1 and correlate positively with (appetitive and aversive) goal values during free-bid trials in study 2. BOLD activity in the right medial OFC ($x = 3, y = 36, z = -18, t = 4.43$) (**a**) and the right DLPFC ($x = 42, y = 42, z = 15, t = 4.29$) (**b**) correlate positively with appetitive goal values and negatively with aversive goal values in a common valuation region.

and the DLPFC were negatively correlated with a trial-by-trial measure of aversive goal values (study 1), that this locus intersects in a conjunction analysis with regions that in a previous study have been found to exhibit positive correlation with appetitive goal values (Plassmann et al., 2007), and that these same areas also correlated positively with overall goal values when subjects need to make decisions about both appetitive and aversive stimuli (study 2). These results suggest that the mOFC and the right DLPFC might encode a goal value signal at the time of choice that applies to both appetitive stimuli and aversive stimuli.

It is important to emphasize that our evidence suggests that the mOFC and the DLPFC play a similar role in computing goal values in appetitive and aversive situations. This does not imply, however, that the exact same network is active in computing both types of valuations. In particular, it does not preclude the likely possibility that these two valuation areas receive differential inputs in the appetitive and aversive cases. In fact, our experimental design also allowed looking for areas that might play a more important role in the computation of appetitive or aversive goal values. We found that activity in the insula correlated more strongly with appetitive goal values than with aversive goal values and that the opposite relationship might hold for the inferior frontal gyrus.

An important question in behavioral neuroscience is whether the goal value circuitry constantly engages in the evaluation of stimuli, even when a choice is not possible, or if it only does so when a choice opportunity is at hand. A recent fMRI study (Lebreton et al., 2009) addressed this question directly and found appetitive goal value signals in the mOFC during both active and passive evaluation situations. In contrast, we did not find significant value activity in this area during the forced trials (although some marginal value coding was found in DLPFC) (Fig. 6). One

possibility is that the presence of the forced bids deactivates an automatic valuation process that would have been present otherwise.

Given the spatial limitation of fMRI, we cannot rule out the possibility that the regions identified in this paper contain neurons that encode separately for the goal value of appetitive and aversive items. An extension of the experimental design of previous studies that have looked at single neuron responses to appetitive goal values (Padoa-Schioppa and Assad, 2008) to the case of aversive items could be used to explore the extent to which different neural populations within OFC and DLPFC encode separately the value of appetitive and aversive items.

With respect to mOFC, our conclusions regarding the existence of a common value region are supported by the literature on the functional anatomy of the OFC, which has shown that this area is uniquely placed as a “hub” that integrates a variety of different inputs such as sensory information from all five senses and visceral and somatosensory information (Carmichael and Price, 1994, 1996). The mOFC also has heavy direct or indirect interconnections with brain areas involved in anticipation, learning and experiencing rewards and punishments, such as the striatum, insular cortex, amygdala, and lateral parts of the OFC, which might also participate in the valuation process. The nature of the inputs supplied by these and other connections is an important open question for future research. More generally, this study contributes to the growing literature on the functions of prefrontal cortex in decision making by showing that the valuation functions of mOFC extend to both the appetitive and aversive realms and thus the notion that this area might act as a general stimulus valuator in the service of decision making.

The paper also contributes to our understanding of DLPFC function. This area is known to carry out a variety of cognitive functions including the top-down modulation of posterior brain regions (Miller, 2000; Miller and Cohen, 2001), maintenance of information in working memory (Levy and Goldman-Rakic, 2000; Curtis and D’Esposito, 2004), and manipulation of information in working memory (Petrides, 2000). Several studies have also shown that the activity of some parts of DLPFC is correlated with various measures of value during perceptual and economic decision-making tasks (Kim and Shadlen, 1999; Wallis and Miller, 2003; Seo et al., 2007). The results in this paper suggest that DLPFC plays a role in the computation of goal values in both appetitive and aversive domains, perhaps by supplying inputs to OFC. Such a hypothesis is supported by anatomical studies in primate brains that have shown that the OFC and the DLPFC are heavily interconnected (Petrides and Pandya, 1999).

References

- Anderson AK, Sobel N (2003) Dissociating intensity from valence as sensory inputs to emotion. *Neuron* 39:581–583.
- Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N (2003) Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci* 6:196–202.
- Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC (2003) Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J Neurosci* 23:9632–9638.
- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Becker GM, DeGroot MH, Marschak J (1964) Measuring utility by a single-response sequential method. *Behav Sci* 9:226–232.
- Bohm P, Linden J, Sonneggard J (1997) Eliciting reservation prices: Becker-DeGroot-Marschak mechanisms vs. markets. *Econ J* 107:1079–1089.
- Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001) Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30:619–639.
- Carmichael ST, Price JL (1994) Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol* 346:366–402.
- Carmichael ST, Price JL (1996) Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol* 371:179–207.
- 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.
- Cloutier J, Heatherton TF, Whalen PJ, Kelley WM (2008) Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *J Cogn Neurosci* 20:941–951.
- Curtis C, D’Esposito M (2004) The effects of prefrontal lesions on working memory performance and theory. *Cogn Affect Behav Neurosci* 4:529–539.
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711.
- Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84:3072–3077.
- Dillon DG, Holmes AJ, Jahn AL, Bogdan R, Wald LL, Pizzagalli DA (2008) Dissociation of neural regions associated with anticipatory versus consummatory phases of incentive processing. *Psychophysiology* 45:36–49.
- Elliott R, Newman JL, Longe OA, Deakin JF (2003) Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23:303–307.
- Elliott R, Agnew Z, Deakin JF (2008) Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. *Eur J Neurosci* 27:2213–2218.
- Elliott R, Agnew Z, Deakin JF (2010) Hedonic and informational functions of the human orbitofrontal cortex. *Cereb Cortex* 20:198–204.
- Gottfried JA, O’Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301:1104–1107.
- Hare TA, O’Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28:5623–5630.
- Hare TA, Camerer CF, Knoepfle DT, Rangel A (2010) Value computations in VMPFC during charitable decision making incorporate input from regions involved in social cognition. *J Neurosci* 30:583–590.
- Horowitz JK (2006) The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods. *Econ Lett* 93:6p.
- Horvitz JC (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96:651–656.
- Jabbi M, Swart M, Keysers C (2007) Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34:1744–1753.
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47:263–291.
- Karni E, Safra Z (1987) “Preference reversal” and the observability of preferences by experimental methods. *Econometrica* 55:11p.
- Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2:176–185.
- Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001) Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* 12:3683–3687.
- Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G (2007) Neural predictors of purchases. *Neuron* 53:147–156.
- Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64:431–439.
- Levy R, Goldman-Rakic PS (2000) Segregation of working memory functions within the dorsolateral prefrontal cortex. *Exp Brain Res* 133:23–32.
- Liu X, Powell DK, Wang H, Gold BT, Corbly CR, Joseph JE (2007) Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. *J Neurosci* 27:4587–4597.

- Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459:837–841.
- McClure SM, Li J, Tomlin D, Cypert KS, Montague LM, Montague PR (2004) Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44:379–387.
- Miller EK (2000) The prefrontal cortex and cognitive control. *Nat Rev* 1:59–65.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102.
- O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003a) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J Neurosci* 23:7931–7939.
- O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ (2003b) Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41:147–155.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Padoa-Schioppa C, Assad JA (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat Neurosci* 11:95–102.
- Pagnoni G, Zink CF, Montague PR, Berns GS (2002) Activity in human ventral striatum locked to errors of reward prediction. *Nat Neurosci* 5:97–98.
- Paulus MP, Frank LR (2003) Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport* 14:1311–1315.
- Petrides M (2000) The role of the mid-dorsolateral prefrontal cortex in working memory. *Exp Brain Res* 133:44–54.
- Petrides M, Pandya DN (1999) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci* 11:1011–1036.
- Peyron R, García-Larrea L, Grégoire MC, Costes N, Convers P, Lavenne F, Mauguière F, Michel D, Laurent B (1999) Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain* 122:1765–1780.
- Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27:9984–9988.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev* 9:545–556.
- Rolls ET, Kringelbach ML, de Araujo IE (2003) Different representations of pleasant and unpleasant odours in the human brain. *Eur J Neurosci* 18:695–703.
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300:1755–1758.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
- Seo H, Barraclough DJ, Lee D (2007) Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb Cortex* 17 [Suppl 1]:i110–i117.
- Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, Dolan R (2005) Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat Neurosci* 8:1234–1240.
- Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007) Differential encoding of losses and gains in the human striatum. *J Neurosci* 27:4826–4831.
- Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T (2003) Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39:701–711.
- Tom SM, Fox CR, Trepel C, Poldrack RA (2007) The neural basis of loss aversion in decision-making under risk. *Science* 315:515–518.
- Ursu S, Carter CS (2005) Outcome representations, counterfactual comparisons and the human orbitofrontal cortex: implications for neuroimaging studies of decision-making. *Brain Res Cogn Brain Res* 23:51–60.
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081.
- Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G (2003) Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron* 40:655–664.
- Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006) Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci* 26:9530–9537.