

# Neurobiology of Value Integration: When Value Impacts Valuation

Soyoung Q Park,<sup>1,2,3</sup> Thorsten Kahnt,<sup>3,4</sup> Jörg Rieskamp,<sup>5</sup> and Hauke R. Heekeren<sup>1,2,3</sup>

<sup>1</sup>Department of Education and Psychology, Freie Universität Berlin, 14195 Berlin, Germany, <sup>2</sup>Neurocognition of Decision Making Group, Max Planck Institute for Human Development, 14195 Berlin, Germany, <sup>3</sup>Berlin School of Mind and Brain, Humboldt Universität zu Berlin, 10117 Berlin, Germany, <sup>4</sup>Bernstein Center for Computational Neuroscience, Charité–Universitätsmedizin Berlin, 10117 Berlin, Germany, and <sup>5</sup>Department of Psychology, University of Basel, 4051 Basel, Switzerland

Everyday choice options have advantages (positive values) and disadvantages (negative values) that need to be integrated into an overall subjective value. For decades, economic models have assumed that when a person evaluates a choice option, different values contribute independently to the overall subjective value of the option. However, human choice behavior often violates this assumption, suggesting interactions between values. To investigate how qualitatively different advantages and disadvantages are integrated into an overall subjective value, we measured the brain activity of human subjects using fMRI while they were accepting or rejecting choice options that were combinations of monetary reward and physical pain. We compared different subjective value models on behavioral and neural data. These models all made similar predictions of choice behavior, suggesting that behavioral data alone are not sufficient to uncover the underlying integration mechanism. Strikingly, a direct model comparison on brain data decisively demonstrated that interactive value integration (where values interact and affect overall valuation) predicts neural activity in value-sensitive brain regions significantly better than the independent mechanism. Furthermore, effective connectivity analyses revealed that value-dependent changes in valuation are associated with modulations in subgenual anterior cingulate cortex–amygdala coupling. These results provide novel insights into the neurobiological underpinnings of human decision making involving the integration of different values.

## Introduction

In everyday life, we choose between options with multiple attributes. The attributes of an option (e.g., shoes) can be qualitatively different (aesthetics and expenses) and are associated with positive or negative values. For successful choice behavior, individuals need to integrate the different values into an overall subjective value.

Behavioral economics has investigated value integration mechanisms to predict choice behavior across a distribution of positive and negative values. Multiattribute utility theory suggests that the subjective value of multiattribute options equals the attributes' weighted sum (Keeney and Raiffa, 1976; Wallenius et al., 2008). Although these models can predict choice behavior well (Huber, 1974; Wallenius et al., 2008), they require that the preference order of one attribute is independent of other attributes. However, human choice often violates this (Keeney and Raiffa, 1976); for example, when selecting a dinner menu with cheese, red wine has a higher value than white wine. But, with fish, white wine has a higher value. Here, an

independent model fails to predict choice, whereas an interactive integration model would successfully predict choice by permitting an extra term for the dependence of attributes.

How the neural systems mediate the value integration is not well understood. The subgenual anterior cingulate cortex (sgACC) has been shown to encode both positive and negative values (Blood et al., 1999; Plassmann et al., 2010). Also, the amygdala represents values independent of valence (Breiter et al., 1996, 2001; Becterra et al., 2001; Gasic et al., 2009). Furthermore, these structures play a key role in both affect (Phelps et al., 2004) and pain regulation (Bingel et al., 2006; Wiech et al., 2008). Thus, the sgACC and the amygdala are ideally suited to facilitate interactive value integration.

In this study, we investigated how the brain integrates values across discrete stimuli into one subjective value to guide decision making. We hypothesized that (1) different values affect each other and (2) the sgACC and amygdala are critically involved. To test these hypotheses, we measured brain activity using fMRI while subjects accepted or rejected offers that were combinations of qualitatively different values of different valence (pain and money). The combination of values included a parametric variation in their intensities.

A well established approach to investigate cognitive processes underlying decision making is to compare cognitive models on behavioral data (O'Doherty et al., 2007; Mazur and Biondi, 2009; Talmi et al., 2009; Bhatt et al., 2010; Navalpakkam et al., 2010). However, if competing models predict the same pattern of choices, behavioral data are limited (Bruni and Sugden, 2007). In these cases, forcing the models to predict neural activity can provide decisive evidence (Glimcher and Rustichini, 2004; Hampton

Received Sept. 21, 2010; revised March 24, 2011; accepted May 4, 2011.

Author contributions: S.Q.P., T.K., J.R., and H.R.H. designed research; S.Q.P. performed research; S.Q.P. and T.K. analyzed data; S.Q.P., T.K., J.R., and H.R.H. wrote the paper.

This work was supported by the Excellence Initiative of the German Federal Ministry of Education and Research (DFG Grants GSC86/1-2009 and EXC 302), the Max Planck Society, and the Swiss National Science Foundation (SNF 100014\_130352).

Correspondence should be addressed to Soyoung Q. Park, Psychology of Emotion/Affective Neuroscience, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: Soyoung.q.park@gmail.com.

DOI:10.1523/JNEUROSCI.4973-10.2011

Copyright © 2011 the authors 0270-6474/11/319307-08\$15.00/0

et al., 2006, 2008; Sanfey et al., 2006; Kable and Glimcher, 2007; Loewenstein et al., 2008). We tested four different subjective value models with either independent or interactive value integration mechanisms. We applied these models directly on behavioral and neural data, looking for decisive information about the implemented mechanism in the brain. Finally, we investigated how different brain regions interact when one attribute's value affects another valuation process.

## Materials and Methods

### Subjects

Twenty-four healthy male subjects (age:  $26.79 \pm 0.66$  years) were included in the study. Subjects reported no psychiatric or neurological disorder. Written informed consent was obtained from all participants after the procedure had been fully explained. The study was approved by the Ethics Committee of the Charité–Universitätsmedizin Berlin.

### Task

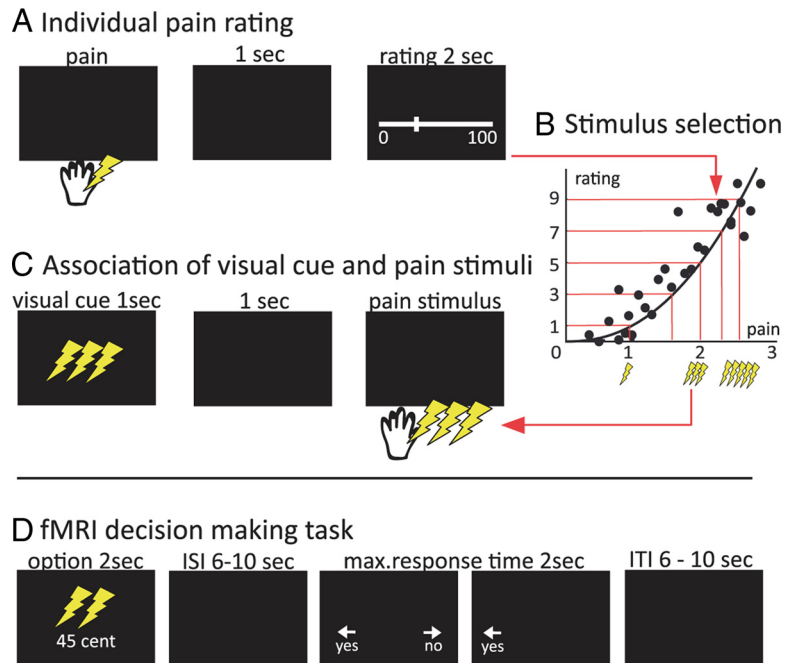
**Individual pain stimulus selection.** Before the scanning session, subjects received 30 mild shocks of varying levels in randomized order and gave ratings on a visual analog scale (VAS) (Price et al., 1994; Brooks et al., 2010). The very left extreme of the VAS was labeled as 0 (not unpleasant at all); the very right extreme was labeled as 100 (worst imaginable unpleasantness) (Fig. 1A). For the tactile-stimulus application, we used a DS5 (Digitimer) stimulator controlled by a stimulation computer. A ring electrode was placed on the back of the left hand between thumb and index finger. For each subject, we fitted a power function to these ratings (Price et al., 1983) and defined five different pain stimuli with equal intervals in subjectively perceived unpleasantness (Fig. 1B). All visual and tactile stimuli as well as response recordings were controlled using Cogent2000 and MATLAB.

**Associating tactile stimuli with visual stimuli.** The set of five different pain stimuli obtained for each individual was then associated with five visual cues via a classical conditioning procedure. In each trial, a visual cue predicting a specific stimulus strength was presented for 1 s. After a 1 s delay, the corresponding tactile stimulus was applied to the subject in paired trials (80%), and no stimulation occurred in unpaired trials (20%). Each association was repeated 10 times (Fig. 1C).

**fMRI decision-making task.** fMRI acquisition consisted of four runs with 60 trials each. In each trial, subjects viewed one offer. The offer was a combination of a monetary amount and a visual pain cue that was learned previously. After a variable delay, subjects either accepted or rejected the offer by a left- or right-hand button press (Fig. 1D). Subjects were told that after the entire experiment, 15 trials of each run would be randomly selected and the accepted offers would be delivered to the subject (both money and pain), whereas the rejected offers would not. We had 12 monetary offers for each subject ranging from 1 to 99 cents (€) (mean 38 cent  $\pm$  2 SEM). For each subject, all combinations of money–pain cue pairs occurred equally often. All pain levels were paired with the complete range of monetary amounts. Before entering the scanner, subjects performed a practice version of the task from which the range of monetary offers was selected for each subject, ensuring that a similar number of offers would be accepted as well as rejected.

### Subjective value models and behavioral analysis

The subjective value models integrated pain and money either independently or they additionally assumed an interaction between both attributes. The interactive term can be thought of as modulating the slope of the value of money as a function of pain. It quantifies by how much the increase in money (i.e., 1 to 99 cents) paired with low pain differs from



**Figure 1.** Multiattribute decision-making task. **A**, Subjects rated tactile stimulations of different strengths on a VAS. **B**, To select five pain stimuli for each individual, we estimated individual power functions using the subjective unpleasantness ratings. **C**, The five selected pain stimuli were then associated with five different visual cues using a classical conditioning procedure. **D**, fMRI experiment. In each trial, subjects saw an offer, which was a combination of a visual pain cue and an amount of money. After a variable delay, subjects either accepted or rejected the offer. Subjects were told that 15 trials would be randomly selected at the end of the experiment and both money and pain would be given in case the selected trial was an accepted offer and that they would receive nothing if it was a rejected offer. ISI, Interstimulus interval; ITI, intertrial interval.

the same monetary increase paired with high pain (Fig. 2A–D). Behavioral studies have suggested nonlinear value functions that allow concavity for positive values and convexity for negative values (Kahneman and Tversky, 1979). For completeness, we modeled the value functions for pain and money in both a linear and nonlinear manner. We refer to these models as (1) linear independent, (2) nonlinear independent, (3) linear interactive, and (4) nonlinear interactive (Fig. 2A–D). Mathematically, all models can be represented as a special case of the nonlinear interactive model, which defines the subjective value of a choice option  $x$  by the subjective value of the monetary amount and the pain level of the option:

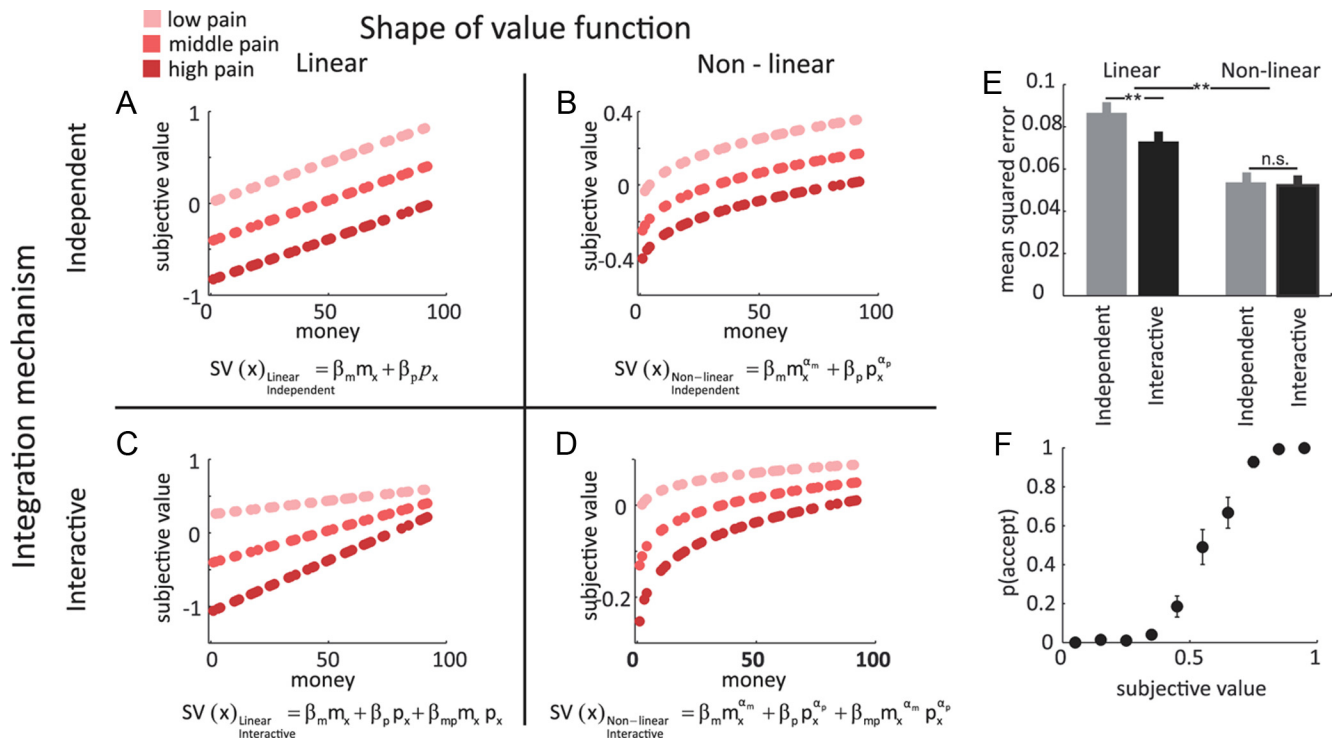
$$SV(x) = \beta_m m_x^{\alpha_m} + \beta_p p_x^{\alpha_p} + \beta_{mp} m_x^{\alpha_m} p_x^{\alpha_p},$$

where  $SV$  is the subjective value,  $m_x$  is the monetary amount,  $p_x$  is the pain level, and the  $\beta$ s represent the weights for money, pain, and the interaction, from left to right. The shape of the value functions for pain and money is modulated by an exponent  $\alpha$  and thus allowed to deviate from linearity ( $\alpha = 1$ ) to be concave ( $\alpha < 1$ ) or convex ( $\alpha > 1$ ). In case the weight  $\beta_{mp}$  for the interaction of pain and money is set to zero, Equation 1 represents the two independent models (Fig. 2A, B), and in case the exponent for the value functions is 1, Equation 1 represents the two linear models (Fig. 2A, C). For all models, we assumed that the probability of accepting an option is a monotonic function of the options' subjective value, as defined by the soft-max choice rule:

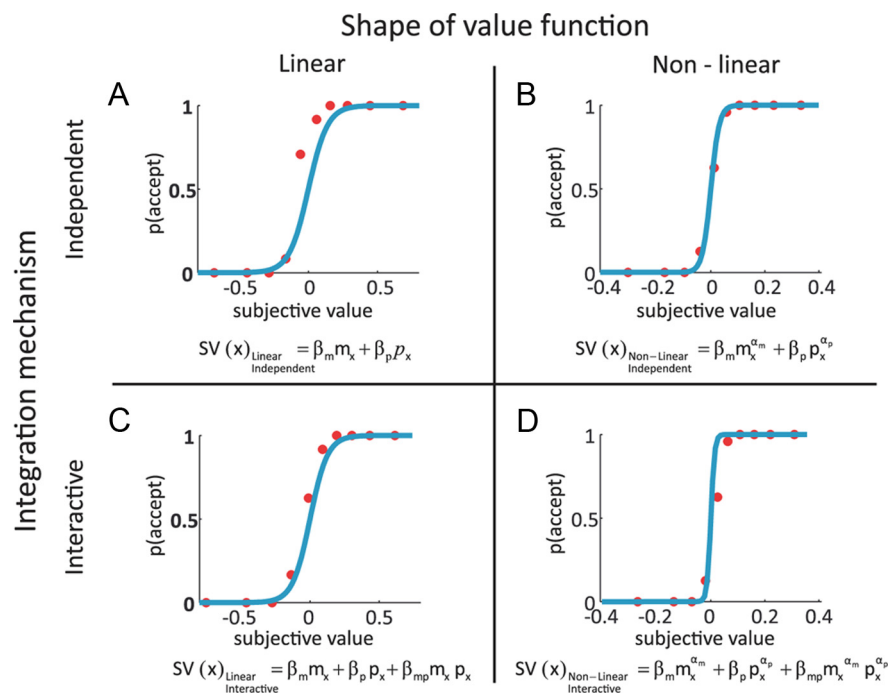
$$p(\text{accept } x) = \frac{1}{1 + e^{-\pi \cdot SV(x)}},$$

where  $\pi$  is a sensitivity parameter defining the slope of the sigmoid function, that is, the choices' stochasticity (the percentage of accepted offers plotted as a function of subjective value of the nonlinear independent model for demonstration; Fig. 2F).

Individual model parameters were estimated using a leave-one-out cross-validation procedure by minimizing the mean squared errors (MSE: average squared difference between the model prediction and subjects' actual choice behavior). Data from three runs were used to fit the free



**Figure 2.** Subjective values of the four models. The subjective values (SV) of the corresponding model as a function of money are plotted. Color of the data points indicates the level of pain (i.e., the darker the points, the higher the pain level). **A**, Linear independent; **B**, nonlinear independent; **C**, linear interactive; **D**, nonlinear interactive. In the interactive models (**C**, **D**), the subjective values from different pain levels are not only shifted on the y-axis but also have different slopes, that is, the difference between high money versus low money is modulated depending on the combined pain level (data from one subject for demonstration). **E**, MSE of the four models in predicting choice behavior. When value functions were modeled in a linear fashion, the interactive models predicted the choice behavior better compared with the independent models ( $t_{(23)} = 4.17, p < 0.0001$ ). However, when value functions were modeled nonlinearly, both integration mechanisms did not differ substantially ( $t_{(23)} = 1.28, p = 0.21$ ). Note that the better predicting model has smaller MSE. Error bars indicate SEM. **F**, The probability to accept increases as a function of subjective value (nonlinear interactive model). Error bars indicate SEM.



**Figure 3.** Predicted and actual percentage of accepted offers. The probability to accept plotted as a function of the subjective values (SV) derived from the four models. Blue lines indicate the model predictions and red dots indicate the actual choice behavior (percentages of accepted offers, binned in 10 categories). Data of a single subject are shown for illustration. **A**, Linear independent; **B**, nonlinear independent; **C**, linear interactive; **D**, nonlinear interactive.

model parameters, and their prediction accuracy was computed on the fourth independent test run. This procedure was repeated four times, each time using a different run as the independent test dataset. The prediction accuracy of a given model was defined as the average MSE in predicting the independent test data across all four cross-validation steps. This procedure allowed us to compare the MSE of four models against each other in predicting choice behavior, independent of the models' complexities (i.e., number of free parameters) (Stone, 1974; Hampton et al., 2008). MSE scores did not significantly deviate from a normal distribution (Kolmogorov–Smirnov test, all  $p$  values  $> 0.7$ ). We thus compared the models' ability to predict choice behavior using a  $2 \times 2$  ANOVA (integration mechanism  $\times$  shape of value function).

*fMRI acquisition and preprocessing*

Functional imaging was conducted on a 3 tesla Siemens Trio scanner with 12-channel head coil. In each of the four runs, 465 T2\*-weighted gradient-echo EPIs containing 33 slices (3 mm thick) separated by a gap of 0.75 mm were acquired. Imaging parameters were as follows: TR = 2000 ms, TE = 30 ms, flip angle = 90°, matrix size =  $64 \times 64$ , and FOV = 192 mm, voxel size =  $3 \times 3 \times 3.75$  mm.

Functional data were analyzed using SPM5 (Wellcome Department of Imaging Neuroscience). The first three volumes of each run were



discarded to allow for magnetic saturation effects. Images were slice time corrected, realigned, spatially normalized to a standard T2\* template of MNI, resampled to 3 mm isotropic voxels, and spatially smoothed using an 8 mm FWHM Gaussian kernel. All included subjects moved less than the size of a single voxel (3 mm; maximal between-scan movement in mm, mean  $\pm$  SEM,  $x = 0.15 \pm 0.02$ ;  $y = 0.36 \pm 0.04$ ;  $z = 0.61 \pm 0.1$ ; in radians, mean  $\pm$  SEM, pitch =  $0.0087 \pm 0.0025$ ; roll =  $0.0031 \pm 0.0004$ ; yaw =  $0.0025 \pm 0.0003$ ).

#### Model-based fMRI data analysis

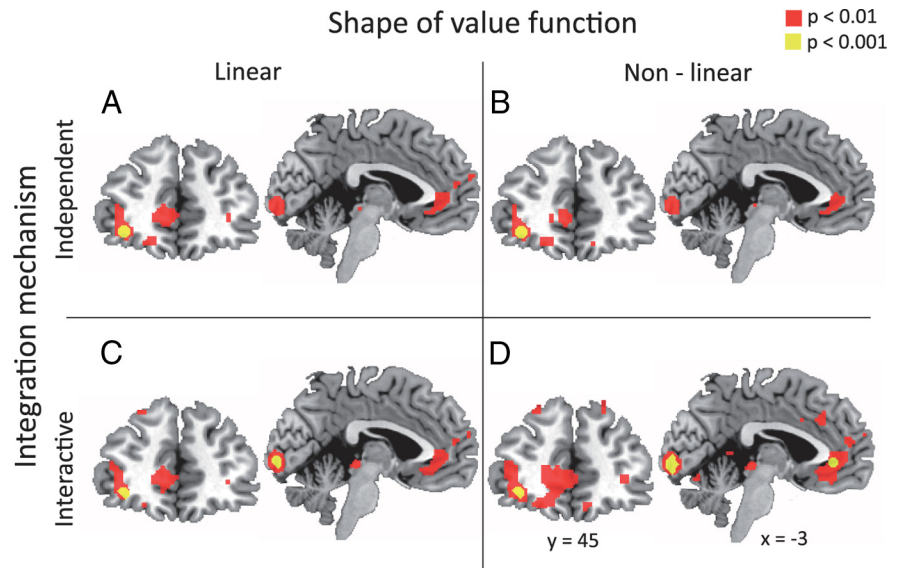
To test the four models against each other at the neural level, for each subject, we set up a GLM with a parametric design (Büchel et al., 1998) for each subjective value model, resulting in four GLMs per subject. Each GLM had three regressors of interest: (1) onset of the offer, (2) the trial-wise subjective value of the offer, and (3) response onset. The subjective value regressor was created by parametrically modulating the stimulus function of the offer onset by the standardized (mean = 0, SD = 1) trial-wise subjective values derived from the four models. The regressors were convolved with a canonical HRF and orthogonalized with respect to the offer onset. All regressors were simultaneously regressed against the BOLD signal in each voxel. The regressors for offer and response onset are identical in all subjective value models; thus, the individual  $t$  maps of the subjective value regressors are proportional to the amount of variance in the BOLD response that is explained solely by each of the subjective value models in each voxel (i.e., effect size). The  $t$  maps of all four models were taken to a second-level random effect analysis.

First, to identify brain regions significantly correlating with the average of all four subjective value models, the four effect size images were averaged and tested ( $p < 0.001$ , uncorrected,  $k = 10$ ). Second, analogous to the behavioral comparison, we set up a second-level random effect analysis, using a  $2 \times 2$  ANOVA. Because the aim of this study was to identify which integration mechanism is used by the brain, the important comparison is the difference between interactive and independent models (and vice versa, collapsing across the value function shapes). The same procedure was applied to compare nonlinear and linear subjective values on the neural level. For this analysis, we applied a threshold of  $p < 0.005$ ,  $k = 5$ , uncorrected within the mask of the average map of subjective values ( $p < 0.05$ ). After having identified the regions in which changes in BOLD signal were significantly better predicted by the interactive models compared with independent models, we extracted the effect sizes of this region. We then performed *post hoc*  $t$  tests to investigate whether the interactive models made better predictions also for both linear and nonlinear value functions separately.

Finally, to resolve the question that could not be answered with the behavioral data, namely the superiority of the integration mechanism within the nonlinear models, we compared the effect sizes of the nonlinear interactive and the nonlinear independent models on a whole brain level using voxelwise paired  $t$  tests ( $p < 0.001$ , uncorrected).

#### Task-dependent changes in connectivity with the sgACC

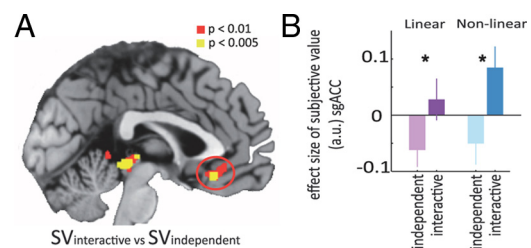
We performed a whole-brain psychophysiological interaction (PPI) analysis (Friston et al., 1997; Kahnt et al., 2009; Park et al., 2010) with the sgACC as a seed region. After having shown that the sgACC is involved in interactive value integration, we aimed to investigate how the interaction between pain and money actually modulates the effective connectivity of the sgACC with any other brain region. In contrast to the standard PPI analysis with only one psychological factor, we set up a PPI using two psychological factors (pain and money). We then searched for changes in effective connectivity with the interaction of pain and money. We first sorted all trials according to their pain and money levels into nine classes (3 money [low, middle, and high]  $\times$  3 pain [low (levels 1 and 2), middle



**Figure 4.** Brain regions encoding the subjective value of the four models. **A**, Linear independent; **B**, nonlinear independent; **C**, linear interactive; **D**, nonlinear interactive. Slices represent coronal (left) and sagittal (right) views of structural brain images with superimposed statistical maps.

**Table 1.** Average effects of subjective values ( $p < 0.001$ ,  $k = 10$ )

		MNI				$t$ value
Region name	BA	$x$	$y$	$z$		
L	OFC	11	-21	33	-15	3.83
	Medial OFC	32	0	39	-3	3.69
L	Central OFC	10	-36	48	-6	4.39
R	dIPFC	8	42	30	51	4.09
L	dIPFC	8	-24	36	54	3.67
L	Parietal cortex	7	-24	-63	54	3.67
R	Parietal cortex	40	-51	-42	57	3.43
L	Occipital lobe	18	-12	-96	6	7.70
R	Occipital lobe	18	27	-99	-6	4.39
L	Cerebellum		-27	-81	-48	3.65
R	Cerebellum		42	-81	-33	4.18

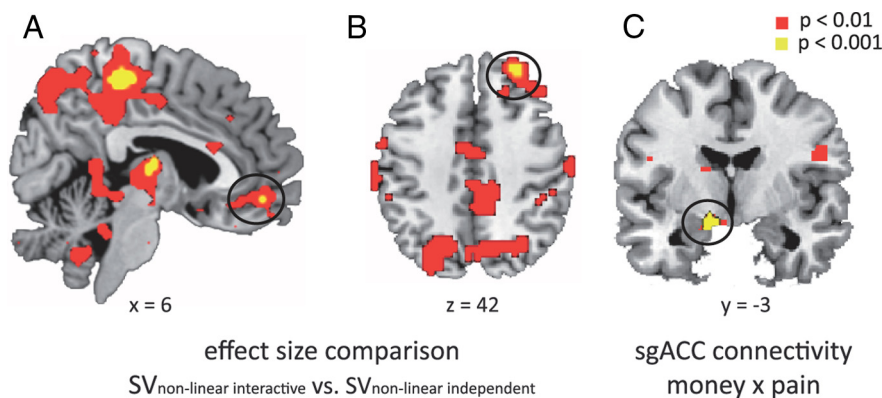


**Figure 5.** sgACC is involved when value affects valuation. **A**, sgACC (0, 27, -15) showing significantly larger effect sizes for the subjective values (SV) of the interactive compared with the independent models in a direct whole-brain model comparison. Slice represents the sagittal view of structural brain image with superimposed statistical map. **B**, This difference was also significant when testing the linear and nonlinear models separately. Error bars indicate SEM.

(level 3), and high (levels 4 and 5)). We extracted the entire time series from each subject in the cluster of the sgACC, in which activity showed significantly higher correlation with subjective values of the interactive models compared with that of the independent models (see Results, Neural representation of different subjective values). We first created regressors for each of the three pain levels. Within each of these three regressors, we coded the three money levels; that is, six TRs following the onset of high money trials were coded as 1, whereas six TRs following the

**Table 2. Comparison between interactive and independent models [collapsed across the curvature of value function;  $p < 0.005$ ,  $k = 5$ , masked with the average map of subjective values ( $p < 0.05$ )]**

Region name	BA	MNI			<i>t</i> value
		<i>x</i>	<i>y</i>	<i>z</i>	
sgACC	25	0	27	−15	2.87
R dlPFC	8	42	33	42	3.43
R OFC	11	12	39	−21	3.23
L Inferior parietal cortex	40	−42	−42	48	2.82
R Inferior parietal cortex	40	45	−42	57	2.82
L Cerebellum	11	−36	−48	−30	3.16
R Thalamus		6	−15	9	3.12
Periaqueductal grey		3	−33	−6	3.45
R Precentral cortex	6	54	−6	54	2.89
L Angular gyrus	7	36	−63	48	2.80
L Middle occipital gyrus	37	−51	−63	20	3.31
L Occipital lobe	19	−27	−66	42	3.10
R Occipital lobe	19	36	−81	21	2.89

**Figure 6.** Whole-brain direct comparison between nonlinear interactive versus nonlinear independent models (**A, B**) and the whole-brain effective connectivity of sgACC (**C**). **A, B**, Brain regions showing significantly larger effect sizes for the nonlinear interactive model compared with the nonlinear independent model. Slices represent sagittal (left) and transversal (right) views of structural brain images with superimposed statistical maps. The circled areas indicate anterior vmPFC [BA 11 (6, 48, −9),  $t_{(92)} = 3.18$ ,  $p < 0.001$ ] (**A**) and dlPFC [BA 9 (30, 45, 42),  $t_{(92)} = 3.78$ ,  $p < 0.001$ ] (**B**). **C**, Functional connectivity between sgACC and left amygdala is enhanced as a function of money offers within the high-pain compared with the low-pain condition [(-12, −3, −12),  $t_{(69)} = 4.14$ ,  $p < 0.001$ ; coronal view of structural brain image with superimposed statistical map].**Table 3. Nonlinear interactive > nonlinear independent ( $p < 0.001$ ,  $k = 10$ )**

Region name	BA	MNI			<i>t</i> value
		<i>x</i>	<i>y</i>	<i>z</i>	
R OFC	10/11	6	48	−9	3.18
R dlPFC	8	30	45	42	3.78
R Posterior cingulate cortex	6	3	−30	57	3.81
L Posterior cingulate cortex	6	−3	−30	72	3.42
R Occipital lobe	19	−33	−78	27	3.77
R Thalamus		12	−12	9	3.57
L Cerebellum		−36	−48	−30	3.48
L Midbrain		−12	−12	−27	3.37
L Parietal cortex	7	−27	−78	45	3.33
R Temporal cortex	37	51	−63	−18	3.32
R Midfrontal cortex	6	36	6	60	3.30
L Precentral gyrus	6	−30	−9	60	3.47
L Parahippocampal gyrus	30	−21	−36	−15	3.25
R Parahippocampal gyrus	30	21	−24	−18	3.23

onset of low money trials were coded as −1, and the middle money trials were coded with zeros. The time window of six TRs was selected to capture the entire hemodynamic response function (Kahnt et al., 2009; Park et al., 2010). These regressors were then multiplied by the normalized time series of sgACC. Thus, for each pain level, the resulting

regressor represents the interaction between sgACC activity and money for one pain level. The middle regressors of pain and money were included in the single-subject model, but were not used to compute the group contrasts because of the smaller number of trials in these classes. Importantly, the created PPI regressors were used as covariates in a separate regression, which also included all the psychological regressors [three onset regressors for three pain levels, each regressor coding the onset of high money (as +1), middle money (as 0), and low money (as −1), convolved with an HRF] and the physiological regressor (the entire time series of sgACC). Because the seed region is defined (the contrast interactive vs independent) in a similar way as the psychological factor of the PPI regressors (interaction between pain and money), the psychological, PPI, and physiological regressors may be correlated. Note that this deviates from the standard PPI approach, in which the seed ROI definition is usually orthogonal to the psychological factor. However, because we entered all regressors simultaneously into the model, the shared variance would not be attributed to any of the regressors. According to this, the PPI regressor explains the incremental variance that is neither explained

by the psychological regressors nor by the physiological regressor. Individual contrast images for sgACC connectivity modulated by money for high-pain and low-pain levels were then entered into second-level *t* tests.

## Results

### Model predictions on choice behavior

The Kolmogorov–Smirnov test revealed that all MSEs did not differ from normal distributions (*p* values for MSE linear independent = 0.996; linear interactive = 0.877; nonlinear independent = 0.993; nonlinear interactive = 0.734). A  $2 \times 2$  ANOVA (integration mechanism  $\times$  shape of value function) with the four models revealed that the nonlinear models had higher predictive power compared with the linear models ( $F_{(1,23)} = 17.17$ ,  $p < 0.001$ ). However, we found a significant interaction effect ( $F_{(1,23)} = 21.35$ ,  $p < 0.001$ ) demonstrating the superiority of the interactive model over the independent model only when the

value functions were modeled linearly. Within the nonlinear value functions, the predictive power of both integration mechanisms did not differ substantially ( $t_{(23)} = 1.28$ ,  $p = 0.21$ ) (Fig. 2E; Fig. 3A–D shows choice behavior of a single subject as a function of the subjective values derived from the four models). Thus, we conclude that, on the basis of choice behavior, it is not possible to identify the integration mechanism. Therefore, we further compared the predictive power of the models directly on neural data to gain insight into the underlying cognitive integration mechanism.

### Model-based fMRI data analysis

#### Neural representation of different subjective values

The four models showed similar patterns of correlation with BOLD responses (Fig. 4). On average, the subjective values of the four models showed significant correlation in the medial prefrontal cortex [mPFC (0, 39, −3),  $t_{(92)} = 3.69$ ,  $p < 0.001$ ], the central orbitofrontal cortex [cOFC (−36, 48, −6),  $t_{(92)} = 4.39$ ,  $p < 0.001$ ], and the dorsolateral prefrontal cortex [dlPFC (−24, 36, 54),  $t_{(92)} = 3.67$ ,  $p < 0.001$ ] (see Table 1 for whole-brain results).

Next, we performed a whole-brain model comparison by statistically testing the effect sizes of the different models. The interactive models showed significantly larger effect sizes in sgACC [BA 25 (0, 27, -15),  $t_{(92)} = 2.87$ ,  $p < 0.005$ ] compared with the independent models (collapsed across linear and nonlinear models) (Fig. 5A, see Table 2 for whole-brain results). In contrast, there were no voxels in which BOLD responses were better predicted by the subjective values of the independent models. Furthermore, we did not find any voxels in which BOLD responses were significantly better predicted by the subjective values of the nonlinear compared with the linear models or vice versa.

A *post hoc* ROI analysis revealed that in the sgACC, the effect sizes of the interactive models were larger than those of the independent models in both linear and nonlinear models separately ( $t_{(23)} = 2.47$ ,  $p < 0.05$ ;  $t_{(23)} = 2.87$ ,  $p < 0.05$ , respectively) (Fig. 5B). Thus, we conclude that values are integrated by means of an interactive integration mechanism and that sgACC plays a crucial role in this function.

#### Identifying the neural integration mechanism

We have shown that both nonlinear models made better behavioral predictions than the two linear models. However, within the nonlinear models, both integration mechanisms (interactive and independent) predicted choice behavior equally well (see above). Therefore, we used the brain data to identify which of the nonlinear integration mechanisms is superior in predicting the BOLD signal in value-sensitive brain regions. Thus, this analysis can serve as a tiebreaker between competing models of brain processes underlying decision making. The nonlinear interactive model revealed voxels with higher effect sizes than the nonlinear independent model in the medial OFC [(6, 48, -9),  $t_{(92)} = 3.18$ ,  $p < 0.001$ ] and the dlPFC [(30, 45, 42),  $t_{(92)} = 3.78$ ,  $p < 0.001$ ] as well as other regions (Fig. 6A,B, Table 3). In contrast, we did not find any voxel in which BOLD changes were significantly better predicted by the nonlinear independent model compared with the nonlinear interactive model. Hence, we conclude that the nonlinear interactive model provides the better description of the neural processes underlying choice behavior among multiattribute options.

#### Effective connectivity of sgACC

Finally, the whole-brain PPI analysis with the sgACC as seed region revealed a significant difference in the money-dependent connectivity modulation when contrasting high versus low pain in the amygdala/sublenticular extended amygdala (SLEA) [(-12, -3, -12),  $t_{(69)} = 4.14$ ,  $p < 0.001$ ] (Fig. 6C).

## Discussion

In the present study, we showed that value affects valuation when advantages and disadvantages are integrated into an overall subjective value. This study provides a concrete example of how neuroimaging directly allows one to test between computational models of decision making and facilitates the evaluation of cognitive computations. Thus, our study supports the promise of neuroeconomics that neuroimaging can significantly contribute to the evaluation of economic questions.

Although independent and interactive subjective value models rely on different assumptions, both models make very similar predictions about the choice behavior. Indeed, in our case, the interactive and independent models performed equally well in predicting subjects' behavior when the value functions were modeled nonlinearly. Thus, behavioral data alone were insuffi-

cient to conclude which integration mechanism (interactive vs independent) best describes the cognitive process of integrating the attributes' values into the subjective value of multiattribute options. Therefore, we went on to compare the predictive power of the models on the neural data. This revealed that interactive models are superior to the independent models in predicting neural activity in the sgACC, independent of the curvature of the value functions. Thus, this analysis of neural activation provides a potential solution for the computational mechanisms of value integration, for which behavioral measures in this study were not informative. Only a few studies so far have compared different models directly on neural data (Hampton et al., 2006, 2008; Montague et al., 2006; Kable and Glimcher, 2007; Rangel et al., 2008). This approach is related to some model-based fMRI studies (Breiter et al., 2001; O'Doherty et al., 2004; Seymour et al., 2004; Kim et al., 2006; Talmi et al., 2009) and provides a concrete example of using neural signals with modeling to identify integration computations that cannot be easily identified by behavioral measures. The identification of the better describing model is essential, even if the two models may yield similar patterns of prediction for choice behavior on one dataset, because this does not imply that this will be the case in other decision situations. Specifically, it is important to identify more accurate descriptions of the underlying process because such models will make new and more precise predictions in future and alternative situations (see also Camerer, 2007). In line with this, future studies should create decision situations in which those models make diverging predictions and compare the models' accuracy in predictions.

Finally, we showed that the connectivity between the sgACC and the amygdala/SLEA was modulated as a function of money only during high-pain conditions. This suggests that interactive value integration relies on the interplay between the sgACC and the amygdala/SLEA.

Subjective values of all four models showed significant correlations with the BOLD signal in medial PFC and OFC. This is in line with evidence suggesting that the ventral part of the medial PFC and OFC encodes the reward value of choice options (Aharon et al., 2001; Daw et al., 2006; Kim et al., 2006; Plassmann et al., 2008; Gasic et al., 2009; Hare et al., 2009; Kahnt et al., 2010, 2011; Philiastides et al., 2010; Smith et al., 2010). Talmi and colleagues (2009) have demonstrated that activity in this region increases with rewards and is attenuated by the prospect of pain. Furthermore, our result that sgACC, together with the amygdala/SLEA, is involved in interactively modulating hedonic experience is consistent with a large body of evidence from cognitive neuroscience. In studies investigating monetary gains and losses, Breiter et al. (2001) have reported that SLEA activity is modulated not only by the prospect of monetary gains and losses but also by their outcomes. In monkeys and rats, analogous regions are involved in regulating fear by exerting inhibitory control over amygdala activity (Sotres-Bayon et al., 2004; Quirk and Beer, 2006; Milad and Rauch, 2007). Studies on fear extinction in rats have shown that stimulating this PFC region modulates amygdala responses, thereby affecting the expression of conditioned responding (Quirk et al., 2003; Rosenkranz et al., 2003). In humans, during extinction and the regulation of learned negative values, the sgACC is actively engaged together with the amygdala independent of the modulation strategy (Phelps et al., 2004; Etkin et al., 2006; Delgado et al., 2008; Schiller and Delgado, 2010). Similarly, sgACC-amygdala coupling is involved in pain regulation such as placebo analgesia and pain habituation (Mayberg et al., 2002; Bingel et al., 2006, 2007). An interesting question is whether our results can be generalized to other types of cost-benefit integra-



tion. Recent evidence suggests distinct valuation subsystems for different types of costs and benefits. For example, Prévost et al. (2010) have shown that delay and effort discounting engage different neural circuits (see also Croxson et al., 2009). Also, besides sgACC and amygdala, ventromedial PFC, and striatum have been shown to play a key role in initial acquisition and modulation of the fear response (Schiller and Delgado, 2010).

In summary, the present study compared different subjective value models with independent and interactive value integration mechanisms directly on fMRI data. This procedure provided neural evidence that an interactive rather than an independent integration mechanism is implemented in the brain. Furthermore, it suggests that the sgACC, in concert with the amygdala, is critically involved in this process. By demonstrating how different values are integrated in the brain, our results substantially extend our knowledge about the neurobiological underpinnings of human choice behavior. Moreover, they contribute to the field of neuroeconomics by showing that direct model comparisons on brain data can be used to uncover cognitive processes and thereby to decide among competing models of decision making.

## References

- Aharon I, Etcoff N, Ariely D, Chabris CF, O'Connor E, Breiter HC (2001) Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32:537–551.
- Becerra L, Breiter HC, Wise R, Gonzalez RG, Borsook D (2001) Reward circuitry activation by noxious thermal stimuli. *Neuron* 32:927–946.
- Bhatt MA, Lohrenz T, Camerer CF, Montague PR (2010) Neural signatures of strategic types in a two-person bargaining game. *Proc Natl Acad Sci U S A* 107:19720–19725.
- Bingel U, Lorenz J, Schoell E, Weiller C, Büchel C (2006) Mechanisms of placebo analgesia: rACC recruitment of a subcortical antinociceptive network. *Pain* 120:8–15.
- Bingel U, Schoell E, Herken W, Büchel C, May A (2007) Habituation to painful stimulation involves the antinociceptive system. *Pain* 131:21–30.
- Blood AJ, Zatorre RJ, Bermudez P, Evans AC (1999) Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci* 2:382–387.
- Breiter HC, Etcoff NL, Whalen PJ, Kennedy WA, Rauch SL, Buckner RL, Strauss MM, Hyman SE, Rosen BR (1996) Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17:875–887.
- 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.
- Brooks AM, Pammi VS, Nussair C, Capra CM, Engelmann JB, Berns GS (2010) From bad to worse: striatal coding of the relative value of painful decisions. *Front Neurosci* 4:176.
- Bruni L, Sugden R (2007) The road not taken: how psychology was removed from economics, and how it might be brought back. *Econ J* 117:146–173.
- Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. *Neuroimage* 8:140–148.
- Camerer CF (2007) Neuroeconomics: using neuroscience to make economic predictions. *Econ J* 117:C26–C42.
- Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009) Effort-based cost-benefit valuation and the human brain. *J Neurosci* 29:4531–4541.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Delgado MR, Nearing KI, Ledoux JE, Phelps EA (2008) Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron* 59:829–838.
- Etkin A, Egner T, Peraza DM, Kandel ER, Hirsch J (2006) Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* 51:871–882.
- Friston KJ, Büchel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6:218–229.
- Gasic GP, Smoller JW, Perlis RH, Sun M, Lee S, Kim BW, Lee MJ, Holt DJ, Blood AJ, Makris N, Kennedy DK, Hoge RD, Calhoun J, Fava M, Gusella JF, Breiter HC (2009) BDNF, relative preference, and reward circuitry responses to emotional communication. *Am J Med Genet B Neuropsychiatr Genet* 150B:762–781.
- Glimcher PW, Rustichini A (2004) Neuroeconomics: the confluence of brain and decision. *Science* 306:447–452.
- Hampton AN, Bossaerts P, O'Doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 26:8360–8367.
- Hampton AN, Bossaerts P, O'Doherty JP (2008) Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc Natl Acad Sci U S A* 105:6741–6746.
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324:646–648.
- Huber GP (1974) Multi-attribute utility models—review of field and field-like studies. *Manage Sci* 20:1393–1402.
- 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—analysis of decision under risk. *Econometrica* 47:263–291.
- Kahnt T, Park SQ, Cohen MX, Beck A, Heinz A, Wrase J (2009) Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *J Cogn Neurosci* 21:1332–1345.
- Kahnt T, Heinzle J, Park SQ, Haynes JD (2010) The neural code of reward anticipation in human orbitofrontal cortex. *Proc Natl Acad Sci U S A* 107:6010–6015.
- Kahnt T, Heinzle J, Park SQ, Haynes JD (2011) Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage* 56:709–715.
- Keeney R, Raiffa H (1976) Decisions with multiple objectives: preferences and value tradeoffs. New York: Wiley.
- Kim H, Shimojo S, O'Doherty JP (2006) Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol* 4:e233.
- Loewenstein G, Rick S, Cohen JD (2008) Neuroeconomics. *Annu Rev Psychol* 59:647–672.
- Mayberg HS, Silva JA, Brannan SK, Tekell JL, Mahurin RK, McGinnis S, Jerabek PA (2002) The functional neuroanatomy of the placebo effect. *Am J Psychiatry* 159:728–737.
- Mazur JE, Biondi DR (2009) Delay-amount tradeoffs in choices by pigeons and rats: hyperbolic versus exponential discounting. *J Exp Anal Behav* 91:197–211.
- Milad MR, Rauch SL (2007) The role of the orbitofrontal cortex in anxiety disorders. *Ann N Y Acad Sci* 1121:546–561.
- Montague PR, King-Casas B, Cohen JD (2006) Imaging valuation models in human choice. *Annu Rev Neurosci* 29:417–448.
- Navalpakkam V, Koch C, Rangel A, Perona P (2010) Optimal reward harvesting in complex perceptual environments. *Proc Natl Acad Sci U S A* 107:5232–5237.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454.
- O'Doherty JP, Hampton A, Kim H (2007) Model-based fMRI and its application to reward learning and decision making. *Ann NY Acad Sci* 1104:35–53.
- Park SQ, Kahnt T, Beck A, Cohen MX, Dolan RJ, Wrase J, Heinz A (2010) Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. *J Neurosci* 30:7749–7753.
- Phelps EA, Delgado MR, Nearing KI, LeDoux JE (2004) Extinction learning in humans: role of the amygdala and vmPFC. *Neuron* 43:897–905.
- Philiastides MG, Biele G, Heekeren HR (2010) A mechanistic account of value computation in the human brain. *Proc Natl Acad Sci U S A* 107:9430–9435.
- Plassmann H, O'Doherty J, Shiv B, Rangel A (2008) Marketing actions can modulate neural representations of experienced pleasantness. *Proc Natl Acad Sci U S A* 105:1050–1054.
- Plassmann H, O'Doherty JP, Rangel A (2010) Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J Neurosci* 30:10799–10808.
- Prévost C, Pessiglione M, Méteureau E, Cléry-Melin ML, Dreher JC (2010)

- Separate valuation subsystems for delay and effort decision costs. *J Neurosci* 30:14080–14090.
- Price DD, McGrath PA, Rafii A, Buckingham B (1983) The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* 17:45–56.
- Price DD, Bush FM, Long S, Harkins SW (1994) A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain* 56:217–226.
- Quirk GJ, Beer JS (2006) Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Curr Opin Neurobiol* 16:723–727.
- Quirk GJ, Likhtik E, Pelletier JG, Paré D (2003) Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *J Neurosci* 23:8800–8807.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Rosenkranz JA, Moore H, Grace AA (2003) The prefrontal cortex regulates lateral amygdala neuronal plasticity and responses to previously conditioned stimuli. *J Neurosci* 23:11054–11064.
- Sanfey AG, Loewenstein G, McClure SM, Cohen JD (2006) Neuroeconomics: cross-currents in research on decision-making. *Trends Cogn Sci* 10:108–116.
- Schiller D, Delgado MR (2010) Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends Cogn Sci* 14:268–276.
- Seymour B, O’Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. *Nature* 429:664–667.
- Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J Neurosci* 30:2490–2495.
- Sotres-Bayon F, Bush DE, LeDoux JE (2004) Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn Mem* 11:525–535.
- Stone M (1974) Cross-validated choice and assessment of statistical predictions. *J R Statist Soc Series B Stat Methodol* 36:111–147.
- Talmi D, Dayan P, Kiebel SJ, Frith CD, Dolan RJ (2009) How humans integrate the prospects of pain and reward during choice. *J Neurosci* 29:14617–14626.
- Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K (2008) Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. *Manage Sci* 54:1336–1349.
- Wiech K, Ploner M, Tracey I (2008) Neurocognitive aspects of pain perception. *Trends Cogn Sci* 12:306–313.