A Behavioral and Neural Evaluation of Prospective Decision-Making under Risk

Mkael Symmonds,[1] **Peter Bossaerts,**[2,3] **and Raymond J. Dolan**[1]

[1]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom, [2]California Institute of Technology, Pasadena, California 91125, and [3]École Polytechnique Fédérale Lausanne, 1015 Lausanne, Switzerland

Making the best choice when faced with a chain of decisions requires a person to judge both anticipated outcomes and future actions. Although economic decision-making models account for both risk and reward in single-choice contexts, there is a dearth of similar knowledge about sequential choice. Classical utility-based models assume that decision-makers select and follow an optimal predetermined strategy, regardless of the particular order in which options are presented. An alternative model involves continuously reevaluating decision utilities, without prescribing a specific future set of choices. Here, using behavioral and functional magnetic resonance imaging (fMRI) data, we studied human subjects in a sequential choice task and use these data to compare alternative decision models of valuation and strategy selection. We provide evidence that subjects adopt a model of reevaluating decision utilities, in which available strategies are continuously updated and combined in assessing action values. We validate this model by using simultaneously acquired fMRI data to show that sequential choice evokes a pattern of neural response consistent with a tracking of anticipated distribution of future reward, as expected in such a model. Thus, brain activity evoked at each decision point reflects the expected mean, variance, and skewness of possible payoffs, consistent with the idea that sequential choice evokes a prospective evaluation of both available strategies and possible outcomes.

Introduction

Evaluating alternative actions is central to decision-making. Many everyday situations require agents to generate a chain of actions (a path through a decision-tree), leading to a distribution of outcomes, which engenders uncertainty. This is a focus in ecology, in which animals forage to ensure intake exceeds minimal need constraints (Stephens et al., 2007), and in finance, in which traders reap bonuses by exceeding a target return from sequential transactions (Panageas and Westerfield, 2009). Common to these examples is that the distribution of possible outcomes (energy, money) differs for each available series of choices.

Decision-making models in finance, psychology, and ecology account for uncertainty (risk) and reward when valuing actions (Markowitz, 1952; Kahneman and Tversky, 1979; Stephens and Charnov, 1982). Growing neural evidence supports the idea that key components of an outcome distribution, such as mean and variance, are explicitly encoded in the brain. However, this literature focuses on immediate returns from single choices (Knutson et al., 2005; Abler et al., 2006; Yacubian et al., 2006; Plassmann et al., 2007; Elliott et al., 2008; De Martino et al., 2009), leaving a relative dearth of knowledge about sequential choice.

In game theory and classical dynamic programming, decision-makers' strategies under every contingency are described by a set of actions that maximize subjective value ("utility"). In sequential choice, these utilities are called "continuation values" because action values are contingent on following a future strategy. Thus, in assigning continuation values, decision-makers must make assumptions about what type of future choices they will make. Standard dynamic programming constrains decision-makers to invoke only optimal choices in the future [optimal continuation value (OCV)]. Critically, an optimal strategy can be planned in advance, implying that "online" updating is irrelevant or irrational (dynamically inconsistent) in the absence of new information (Dekel et al., 1998; Epstein and Schneider, 2003). An OCV decision-maker, when presented with decision 1 followed by decision 2, makes the same set of choices even if the order is reversed (as long as no new information is presented until after choices are made).

However, decision-makers can assign utilities to options assuming that they might not take the optimal choice in the future. This might occur if choices became unexpectedly constrained, when planned strategies would no longer be available. All available strategies (rather than just the preplanned "optimal" strategy) are taken into account before each choice, for example, by assuming that future choices are distributed randomly [average continuation value (ACV)]. This entails that strategies are dynamically reevaluated and action values recalculated depending on which strategies are available. This scenario allows for dynamic inconsistency, in which future choices can depend on the order in which options are presented (Machina, 1989).
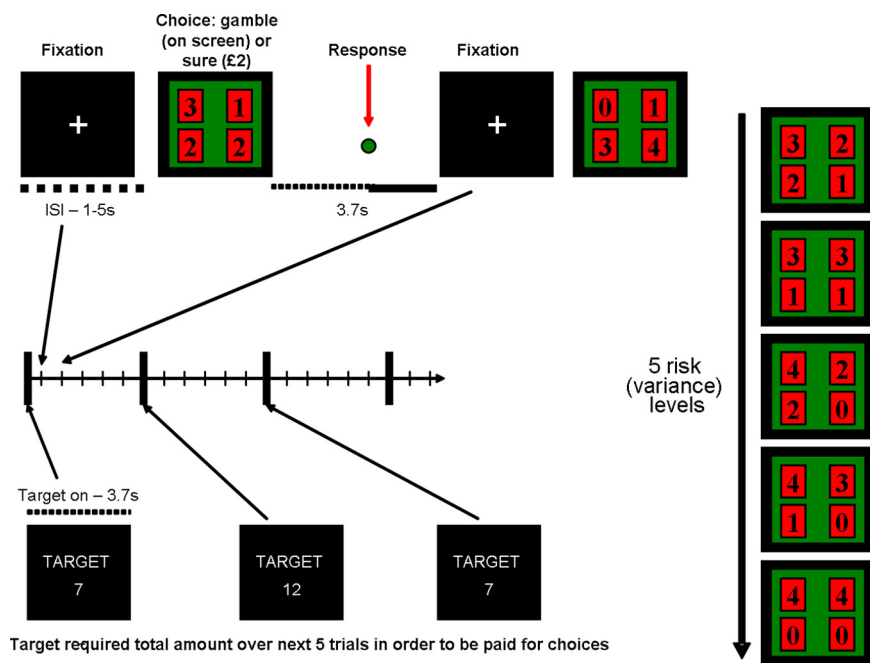
Figure 1. Task structure and timings. Trials were grouped into blocks of five. To commence each block, a financial target appeared on the screen for 3.7 s (4 levels for behavioral experiment outside scanner: 5, 9, 11, 13; 2 levels inside scanner: 7, 12). A fixation cross was shown at the beginning of each trial for 1–5 s (jittered). A lottery then appeared on the screen, and subjects had up to 3.7 s to indicate by button press their choice (gamble or £2 for sure). Lotteries were represented by four red cards, with numbers (0, 1, 2, 3, or 4) indicating monetary values in pounds, each with a probability of 0.25 of being selected. A computerized random-number generator was used for selection; subjects were shown the selected trial and block outcomes. Total earnings were the summed total return from the 10 selected trials. ISI, Interstimulus interval.

Thus, sequential decision-making poses two clear problems. First, do humans evaluate the distribution of outcomes when planning choice? Second, do individuals assume optimal future choices when making sequential decisions? Here, we tested different models of strategy valuation and planning, simultaneously acquiring neural data [using functional magnetic resonance imaging (fMRI)]. We hypothesized that neural activity evoked in single-shot decision paradigms also supports decision variables mediating sequential choice.

Materials and Methods

Behavioral experiment

The study was approved by the Institute of Neurology (University College London, London, UK) Ethics Committee. Seventeen subjects (age range, 22–36 years; seven male) participated; one dropped out from scanning because of claustrophobia and was excluded from analysis. Monetary earnings were between £18 and £28, including a fixed £10 participation fee. Stimuli were presented on a standard personal computer using Cogent presentation software (Wellcome Trust Centre for Neuroimaging, London, UK) run in MATLAB (version 6.5; Math-Works). Choices were made by key-press selections on the computer keyboard of a standard personal computer.

We provided instructions with a 15 min verbal tutorial, to ensure that subjects understood the paradigm. In each block of the task, subjects were required to make five sequential choices between a sure or risky alternative. For each trial, a lottery was represented on screen using a picture of four cards (Fig. 1), and subjects indicated by a button press their choice to gamble or selection of a fixed sure amount of £2. Different numbers on the cards (0, 1, 2, 3, 4) indicated monetary value in pounds. We used five different card combinations, generating five lotteries with matched expected value of £2 but with different variance. All lotteries had symmetrical outcome distributions (i.e., were unskewed), with a mean return of £2. On each block, the same five lotteries were presented once, using a randomized order of presentation, necessary if one is to detect dynamically inconsistent choices. Critically, no feedback about out-

comes was given during the task. This constraint allowed us to focus on a situation in which individuals do not need to adjust their strategies in response to feedback, enabling us to distinguish whether individuals adhere to a predetermined strategy, regardless of the particular sequence in which the options are presented (consistent with classical dynamic programming models in which choices are made based on the optimal continuation value), or whether they continuously reevaluate, taking into account a range of available strategies on each trial.

We altered the distribution of possible outcomes and the relative utility of both the gamble and sure option by imposing financial targets (on each block). At the beginning of each block, a target appeared on the screen (four levels used outside scanner: 5, 9, 11, 13; two levels used inside scanner: 7, 12). Subjects were told that 10 trials would be randomly selected at the end of the experiment from all sessions (inside and outside scanner) in which all trials had an equal chance of being picked. For selected trials, if the required target had been reached in that block, the outcome of that trial would be paid out (i.e., 10 trials are chosen to determine pay, contingent on whether the target was reached at the end of each block). This would be £2 if they had picked the certain fixed amount or whatever the outcome of the lottery (determined by a random selection of one of the four cards), if they had chosen to gamble. If the target had not been reached, then no money would be won from the trial regardless of choice.

Subjects were instructed to try to win as much money as possible, remembering that the total amount won would depend on reaching the target and whether they chose to gamble or not. Subjects were informed that outcomes would be recorded for all their choices in the experiment but that these would not be shown at the time. Subjects were not explicitly told that only five types of gamble would be shown. However, to acquaint them with the task and to engender the idea that actual outcomes were recorded from each choice made, five practice blocks using identical gambles were run with full feedback.

Analysis. We initially categorized trials by two factors, current target level and the variance (risk) of lottery presented. These data were assessed by ANOVA and multiple regression implemented in SPSS (SPSS for Windows, release 12.0.1, 2001; SPSS Inc.). We then analyzed choices by block. There are $2^5 = 32$ possible combinations of choices in each block, and we refer to each of these combinations or trajectories of choices as a strategy, denoting strategy n as s_n ($n = 1, \ldots, N$). The frequency with which each strategy was chosen was compared with simulated strategy choice frequencies by χ^2 test. We simulated blockwise choice frequencies using mechanistic binomial choice models (described below) and estimated the best-fitting parameters of these models using an application of the method of simulated moments (McFadden, 1989). This estimation is based on comparing observed frequencies of choices with simulated frequencies, derived from an underlying structural model. Free parameters were optimized with a nonlinear simplex search algorithm in MATLAB 7.0. We selected the best-fitting model and assessed relative model performances by a comparison of criterion values on an individual subject basis.

Behavioral modeling. Our behavioral models have two components. The first component provides a model of valuation, calculating an expected utility per strategy (V_n, $n = 1, \ldots, N$) given the distribution of outcomes a particular strategy generates. In other words, we apply a utility function to each distribution of outcomes, to generate a single number representing the subjective value of each strategy. The second component models how future choice is incorporated, in which we im-

plement three separate continuation value models. These models specify how strategies are compared and which strategies influence the value of the current choice (to gamble or not to gamble). On the first trial in a block, there are 16 possible strategies given a choice to gamble and 16 possible strategies given a choice of the sure amount. This reflects the fact that there are five ordered binary choices between a lottery and a certain payout, so 32 (2^5) possible sets of choices are available in any given block. Each of these strategies has its own outcome distribution. Note that the set of available strategies reduces as sequential choices are made such that, by the fifth trial in a block, only two possible alternative strategies are available (to gamble or opt for the sure outcome). In other words, the set of possible strategies at any given trial is contingent on previous choices in the block and the order in which the gamble options have been presented.

Valuation model. We used a mean–variance–skewness (MVS) model in which the distribution of outcomes is evaluated by a weighted linear sum of its statistical moments. We assign values (utilities), $V_{s,t}^h$, to available strategies s, on each trial t, for every target level h, and calculate utility given the predicted distribution of outcomes resultant from each strategy s. The set of strategies evaluated are contingent on previous choices in each block. For example, on trial $t = 4$, there will be four possible available strategies to evaluate, given a certain sequence of simulated or actual choices for trials $t = 1,2,3$. Each of these strategies will generate their own distribution of possible numerical outcomes. The probability distribution of outcomes for each strategy will alter depending on target level h.

Let $\mathbf{B}_{n,t}^h$ comprise the set of discrete outcomes given strategy s_n on trial t, where $\mathbf{B}_{n,t}^h(j)$ indexes the jth outcome from this set, and $P_j(\mathbf{B}_{n,t}^h(j))$ indicates the probability of the jth outcome.

In this formulation, strategy value on trial t is specified as follows:

$$V_{s,t} = \mathrm{E}(X) - \rho \cdot \mathrm{Var}(X) + \lambda \cdot \mathrm{Skw}(X).$$

$E(X)$ denotes the expected, (mean) value of the distribution of outcomes from strategy s_n:

$$\mathrm{E}(X) = \mathrm{E}(\mathbf{B}_{n,t}^h) = \sum_j P_j(\mathbf{B}_{n,t}^h(j)) \cdot \mathbf{B}_{n,t}^h(j).$$

$\mathrm{Var}(X)$ denotes variance of outcomes:

$$\mathrm{Var}(X) = \sum_j P_j(\mathbf{B}_{n,t}^h(j)) \cdot \langle \mathbf{B}_{n,t}^h(j) - \mathrm{E}(X)\rangle^2,$$

and $\mathrm{Skw}(X)$ denotes skewness of outcomes:

$$\mathrm{Skw}(X) = \sum_j P_j(\mathbf{B}_{n,t}^h(j)) \cdot \langle \mathbf{B}_{n,t}^h(j) - \mathrm{E}(X)\rangle^3.$$

ρ is a coefficient reflecting aversion to variance in outcomes, and λ reflects the degree of positive skew seeking behavior. Using a free parameter for skewness allows modeling of a wide range of preferences (skewness is a second-order approximation of risk), because this model accounts for preferences for relative losses and gains independently of the spread of outcomes.

We present the results using one valuation model (mean–variance–skewness) in the main text. We also tested alternative utility models, although these models are not directly statistically comparable because they are non-nested with a different number of parameters (see supplemental data, available at www.jneurosci.org as supplemental material). We did not implement more complex variants of prospect theory with differential weighting of relative losses and gains because an MVS model incorporates some aspects of this behavior while being much easier to fit (because it is linear in its arguments).

Continuation value models

To simulate choice, our continuation value models perform a tree search of all possible choice (action) and outcome (state) combinations from current trial t to the end of each simulated block. This search is contingent on (i.e., constrained by) previous choices. We recalculate the value of available strategies on each trial, and, as the block proceeds, the num-

ber of possible strategies available reduces such that by trial t there will be 2^{6-t} possible strategies remaining.

Optimal continuation value. This model assumes that agents pick the choice combination (trajectory) of all possible alternatives that maximizes utility by the end of each block. In other words, the decision-maker compares the current options on each trial (select the sure amount or to gamble), and only the trajectories giving a predicted outcome distribution that maximizes utility determines choice. Thus, OCV prescribes the choices for a prospective decision-maker who acts in accordance with classical dynamic programming principles. Thus, in accordance with standard dynamic programming models, OCV is oblivious to the order in which options are presented.

We assign action values (Q) to the binary options gamble (specified as Q^1) or sure (specified as Q^0), calculated on every trial, as the decision-maker progresses through the decision tree. These two action values are then compared with determine choice:

$$Q^0 = V_{\bar{n}}, \quad \text{where } \bar{n} = \arg\max_n\{V_n \mid \text{sure choice}\},$$

$$Q^1 = V_{\bar{n}}, \quad \text{where } \bar{n} = \arg\max_n\{V_n \mid \text{gamble choice}\}.$$

n indexes the possible continuation trajectories (available strategies) or branches of the decision tree.

Note that the OCV will remain the same on every trial within a block when a subject adheres to the strategy selected at block outset. In the case of a deviation from an optimal trajectory, the next best (utility maximizing) strategy is taken from the remaining options available. Importantly, in the case of these deviations, the OCV model prescribes that appropriate correction is taken based on always trying to follow the utility-maximizing strategy.

Average continuation value. This model entails that agents calculate the average value or utility of each of the two alternative choices on each trial (i.e., choosing to gamble or take the sure amount) rather than forecasting with respect to optimal continuation trajectories. This model does not require that agents have an explicit plan of future choices and is akin to a model in which the current choice is made under an assumption that choices are made randomly for the rest of the block. Every possible strategy influences current continuation values. As such, the decision-maker can be thought of as myopic:

$$Q^0 = \left\{\frac{1}{n}\sum_{n=1}^{\max n} V_n \middle| \text{sure choice}\right\},$$

$$Q^1 = \left\{\frac{1}{n}\sum_{n=1}^{\max n} V_n \middle| \text{gamble choice}\right\}.$$

Sure continuation value. This model assumes that agents weigh up the current choice against a benchmark of taking the sure option for the remainder of the block. This implements a simple heuristic in which the choice to gamble or take the sure amount is made given a fixed benchmark:

$$Q^0 = V_{\bar{n}}, \quad \text{where } \bar{n} = \{V_n \mid \text{sure choice for } T = t, \ldots, 5\},$$

$$Q^1 = V_{\bar{n}}, \quad \text{where } \bar{n} = \{V_n \mid \text{gamble choice for } T, \text{ sure choice for}$$

$$T = t + 1, \ldots, 5\},$$

where t indexes the current trial (T) in the block ($t = 1, \ldots, 5$).

For a given target, several strategies can lead to similar distributions of outcomes. However, strategies will differ in their outcomes depending on the target level. Moreover, a critical feature for all these decision rules is that subjects' previous choices within a block determine the remaining available strategies to be evaluated. These models assume that the full space of possible actions and outcomes is known. We make this simplifying assumption to render model estimation tractable (i.e., specifically we do not incorporate uncertainty about future options). This is not unreasonable given that the task has a simple repeating structure with the

same five lotteries being presented on each block throughout practice, behavioral, and scanning sessions.

Numerical example. We provide a simple numerical example of how these models work in practice (supplemental Fig. S1, available at www. jneurosci.org as supplemental material). Imagine that you are faced with a two-stage sequential decision between a gamble (g) and a sure amount (s) of money, which is fixed at £2. The first decision (decision X) is whether to accept a 50:50 gamble giving either £4 or £1. The second decision (decision Y) is whether to accept a gamble giving a 75:25% chance of winning £3 or £0, or again opting for a sure amount of £2. There are four possible strategies to consider (ss, sg, gs, gg combinations, which we refer to as strategies A, B, C, D), each giving a different distribution of outcomes. In our models, these distributions are evaluated according to a utility function (U), to give four separate numbers, or utilities, one per strategy.

For example, we now assign numbers to these utilities for illustrative purposes: U(ss) = 10; U(sg) = 8; U(gs) = 4; U(gg) = 7. Imagine now having to choose a current action. In our model, if you were an optimal decision-maker, you would compare the highest utilities given each choice [in this case, U(ss) for a sure choice, U(gg) for a gamble choice]. Because U(ss) > U(sg) > U(gg) > U(gs), you prefer to make a sure choice on the current trial. For the next decision, you again make a sure choice, now comparing U(ss) = 10 versus U(sg) = 8. Thus, you have selected strategy A. What if you are a decision-maker conforming to an average continuation value model? In this case, you weigh up the average utility of outcomes from each current choice [i.e., you compare (U(ss) + U(sg))/2 = 9 with (U(gs) + U(gg))/2 = 5.5]. In this example, you also prefer to make a sure choice [as U(s, ·) > U(g, ·)]. For the subsequent decision, you compare U(ss) = 10 and U(sg) = 8 and again make a sure choice, following strategy A.

Now let us consider a situation in which the order of decisions is reversed (supplemental Fig. S2, available at www.jneurosci.org as supplemental material). There remain four strategies, but their order has changed such that we have the following: U(ss) = 10; U(sg) = 4; U(gs) = 8; U(gg) = 7. If you are an optimal continuation value decision-maker, you choose the sure option as U(ss) > U(sg) > U(gg) > U(gs), followed by another sure option [as U(ss) > U(sg)]. The order has no effect on the ranking of the strategies, and you make a dynamically consistent choice by following strategy A again. If you are an average continuation value decision-maker, you will compare (U(ss) + U(sg))/2 = 7 with (U(gs) + U(gg))/2 = 7.5 and pick the gamble choice initially. On the next decision, you make a sure choice, as U(gs) > U(gg), such that now you follow strategy B and have made a dynamically inconsistent choice as the order in which the options are presented has affected choices. A sure continuation value (SCV) decision-maker would make a consistent choice in this example, in which the "sure, sure" strategy A has the highest utility but also can make dynamically inconsistent choices if this is not the case. The actions actually selected (and whether they will be dynamically consistent) will depend in practice on the specific utilities assigned to the available choices by the decision-maker. Note that, in these models, order independence (dynamic consistency) only holds if there is no new information arriving, which is the case in this experiment in which no trial-by-trial feedback is given, also that these models reflect different methods of valuation and planning (i.e., an anticipated selection of choices) rather than testing execution of a preformed plan (i.e., self-consistency).

Action selection. We account for randomness in choice by the addition of noise at action selection (modeled by a logistic choice function). Thus, the predicted probability of choosing to gamble on a given trial is given by the following:

$$P(\text{gamble choice}) = \frac{1}{1 + e^{-\frac{1}{\sigma}(Q^1 - Q^0)}},$$

where σ is a free parameter.

Model estimation. We based our model estimation on a comparison of the observed frequencies of block-by-block choices (i.e., strategies) with simulated frequencies, derived from each of the underlying structural models outlined above. The models generate a choice per trial per simulated block (using the probabilistic action selection rule), from which we

calculate the simulated frequency (φ) with which each strategy is chosen. We ran 1200 simulated blocks per model, across all six target levels, with a randomized trial order per block. These simulated frequencies are then compared with actual observed choices (z), using the method of simulated moments (McFadden, 1989). $z(i)$ is a vector of choices over available strategies on block i, with its elements taking the value 1 for the chosen strategy and 0 otherwise:

$$y_i = z(i) - \phi,$$

$$D = \min_{\theta} \left[\frac{1}{N} \sum_{i=1}^{N} y_i \right]' \Omega^{-1} \left[\frac{1}{N} \sum_{i=1}^{N} y_i \right],$$

where y_i is a vector of observations from one block i (observed − simulated frequencies), Ω is a weighting matrix, and D, the criterion function, is the weighted sum-of-squares difference between the observed and simulated frequencies across all blocks ($i = 1, \ldots, N$). Optimization of D (which finds the best-fitting set of parameters θ) is performed in a two-step procedure. Initial unweighted estimates are derived with $\Omega = \mathbf{I}$ (identity matrix). A weighted optimization is then performed. To estimate the precision of the observations, we calculate the covariance matrix (Ω) of the differences between simulated and observed frequencies that come out of an unweighted optimization:

$$\Omega = \frac{1}{N} \sum_{i=1}^{N} \left[y_i - \frac{1}{N} \sum_{i=1}^{N} y_i \right] \left[y_i - \frac{1}{N} \sum_{i=1}^{N} y_i \right]'.$$

To make Ω invertible, it is necessary to aggregate unchosen strategies, otherwise the weighting matrix is rank deficient. The estimated precision is then the inverse of this covariance matrix (Ω^{-1}). We weight observations according to this precision in performing the weighted optimization to calculate an unbiased estimator. This method of moments criterion function D is not differentiable in the parameters (there are step changes in the value of the function as the parameters vary); hence, we use a simplex search method to optimize the parameters with respect to the criterion function D (by using the Nelder–Mead simplex algorithm implemented in MATLAB). We use the method of simulated moments to optimize the models because the problem of multinomial sequential choice is high dimensional and computationally difficult to integrate. This means that we cannot use Bayesian methods to get a measure such as the Bayesian information criterion. In these circumstances, the method of simulated moments provides a robust way of optimizing models, and the criterion function acts as a likelihood estimate that allows comparison of our model space.

The optimized criterion value D is a direct measure of the residual sum of squared error of each model. D (multiplied by the number of observations) is χ^2 distributed (Hansen's J statistic) (Hansen, 1982). Relative model likelihoods calculated from χ^2 statistics are not comparable for non-nested models. However, to the extent that the number of parameters are equal (for a given utility and noisy choice model), criterion values can be directly compared. Hence, inverse criterion values (D^{-1}), reflecting relative goodness-of-fit, were directly compared for best-fitting models on an individual subject basis. All expected utility and prospect theory models have two free parameters (σ and ρ), whereas MVS models have three free parameters (ρ, σ, and λ).

Functional MRI

All subjects had previously completed the behavioral experiment and understood that the task structure and presented lotteries were identical. We used two target levels during scanning (7 and 12). Visual cues were projected onto a screen, visible via an angled mirror mounted on the MRI head coil. Choices were indicated by pressing a button box with the right index finger, and responses were recorded using Cogent presentation software.

Scanning parameters. We acquired gradient echo T2*-weighted echo-planar images (EPI) with blood oxygen level-dependent (BOLD) contrast, on a 3 T head scanner (Magnetom Allegra; Siemens Medical). Imaging parameters were as follows: 48 oblique transverse slices; slice thickness, 2 mm; gap between slices, 1 mm; repetition time, 3.1 s; echo

time (TE), 30 ms; field of view, 192 × 192 mm². We used an EPI sequence that optimized for BOLD sensitivity in the orbitofrontal cortex (OFC) using a combination of an increased spatial resolution in the readout direction and a reduced echo time (Weiskopf et al., 2007). Together with the oblique orientation of the slice acquisition, this can compensate and recover for potential signal loss in OFC, one of our regions of interest. During the same experimental session, a T1-weighted image was obtained for anatomical reference. To correct for geometric distortions induced in the EPIs at high field strength, we collected field maps based on dual echo-time images (TE1, 29 ms; TE2, 19 ms) and processed these using the statistical parametric mapping SPM5 field-map toolbox (Hutton et al., 2002) to produce a voxel displacement map indicating the field distortions.

Images were realigned with the first volume, normalized to a standard EPI template, and smoothed using an 8 mm full-width at half-maximum Gaussian kernel. Unwarping was performed using the routine in SPM5, correcting for distortions in each acquired image by combining the measured field maps with estimated susceptibility-induced changes attributable motion. Realignment parameters were inspected visually to identify any potential subjects with excessive head movement. Data were analyzed in an event-related manner using a general linear model, with the onsets of each stimulus modeled as a δ function. To capture all variance of interest (i.e., the modulation of neural response preceding each choice), δ functions were placed halfway between the onset of the presentation screen and the subsequent key-press response. Regressors of interest (see Results) were generated by convolving the stimulus functions with a hemodynamic response function. First-order temporal derivatives of each of these convolved functions were included to ensure that any neural activity related to cognitive processes of interest within an approximately ±2s window period should be captured by the convolved δ functions placed at the halfway point (Friston et al., 1998). This also avoids the need to constrain the model by making predictions concerning the timing of the neural responses to the different regressors.

Our contrasts of interest purely concern responses parametrically modulated by specific stimulus dimensions, reflecting activity independent of the regressors modeling nonspecific responses to stimulus presentation. Covariates of no interest comprised the onsets of the target screens and subject-specific realignment parameters from the image preprocessing to account for motion-related artifacts in the images that were not eliminated in rigid-body motion correction. BOLD data from blocks in which a response had been missed were factored out by explicitly including a regressor for these error trials. All data were analyzed using statistical parametric mapping software (SPM5; Wellcome Trust Centre for Neuroimaging). Trial-type-specific β values of linear contrasts were estimated, and these were entered into t tests using random-effects analysis to provide group statistics.

Presentation of data and images. Figures are constructed by thresholding second-level SPM t images at $p < 0.005$, and superimposing data on a mean image across all participants. Stereotactic coordinates are reported in Montreal Neurological Institute (MNI) space (Mazziotta, 2001). For the contrasts of interest, results are reported at a threshold of $p \leq 0.001$ uncorrected. We also report results with small-volume correction for regions of interest dictated by previous studies at $p < 0.05$ (a 6 mm radius sphere centered on a priori coordinates) (for details, see supplemental Tables S6, S7, available at www.jneurosci.org as supplemental material).

Results

Behavioral

Trial-by-trial choices

We first analyzed subjects' choices in terms of a decision to gamble or opt for the sure amount, on a trial-by-trial basis, across all sessions (inside and outside scanner). We observed a linear relationship between risk (variance) of an individual gamble and the percentage of time that subjects chose the gamble over the sure alternative (Fig. 2). A repeated-measures ANOVA demonstrated a significant main effect of both riskiness of each gamble [$F_{(2.86,42.95)} = 2.88$, $p = 0.049$ (Greenhouse–Geisser corrected de-
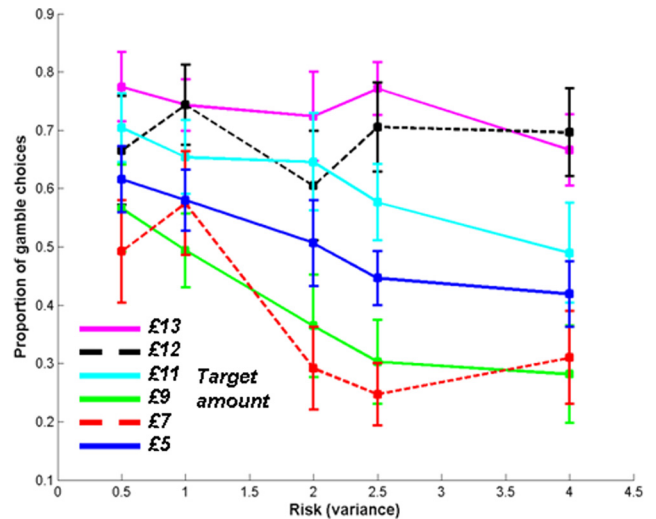

Figure 2. Analysis of trial-by-trial choices. Graph showing the proportion of time that the gamble option was selected overall, for each of the five presented lotteries [abscissa is the variance of each individual lottery in (pounds)²]. Data plotted for both choices in the behavioral experiment outside the scanner (solid lines: using target amounts £5, £9, £11, £13) and choice made inside the scanner (dashed lines: using target amounts £7, £12). Error bars show SE ($n = 16$). The proportion of gamble choices linearly decreased with increasing risk of each lottery, demonstrating increasing risk aversion ($p = 0.049$). There was also a significant linear interaction between the risk of each presented lottery and the target level ($p = 0.007$).

grees of freedom, $\varepsilon = 0.72$); Mauchly's test for sphericity: $\chi^2_{(9)} = 19.49$, $p < 0.05$] and target level [$F_{(5,75)} = 16.32$, $p < 0.001$ (within-subjects contrasts; negative linear effect of risk: $F_{(1,15)} = 9.64$, $p = 0.007$, $r = 0.63$; linear effect of target: $F_{(1,15)} = 42.37$, $p < 0.001$, $r = 0.86$)] (Fig. 2). There was also a significant interaction between risk and target level ($F_{(20,300)} = 2.72$, $p < 0.001$) such that, at higher target levels, the slope of the linear relationship was reduced. There was no tendency for subjects to be more risk seeking at the beginning or end of the blocks, with neither a linear nor quadratic effect of time point within a block on the probability of choosing to gamble (risk × target level × time ANOVA; linear contrast: $\beta = -0.023$, $r^2 = 0.001$, $p = 0.679$; quadratic contrast: $p = 0.42$).

Analysis of choices by block

Descriptively, subjects switched strategy in a systematic manner as the target level changed (Fig. 3). For low target, subjects tended to choose strategies involving fewer lottery gambles. However, for higher targets, as the chance of getting nothing increased, subjects chose strategies involving more lottery gambles, thereby increasing expected return. There was considerable heterogeneity in strategy selection, particularly for medium target levels. Analyzing group data across all subjects demonstrated that choices were significantly different from random (χ^2 test against random choice, df = 155, $p < 0.001$).

A comparison of each of the decision-making models is illustrated in Figure 4A (for details of modeling analysis and alternative utility models, see Materials and Methods and supplemental text available at www.jneurosci.org as supplemental material). In absolute terms, a model of average continuation value obtained the lowest optimized weighted criterion value, or difference between predicted and actual choice frequencies (ACV: mean ± SEM, $D = 6.3 \pm 1.5$; OCV: mean ± SEM, $D = 11.9 \pm 3.0$; SCV: mean ± SEM, $D = 29.5 \pm 7.4$). We also compared models with random choice (in which all strategies are selected with equal frequency) to give an absolute measure of accuracy. A random
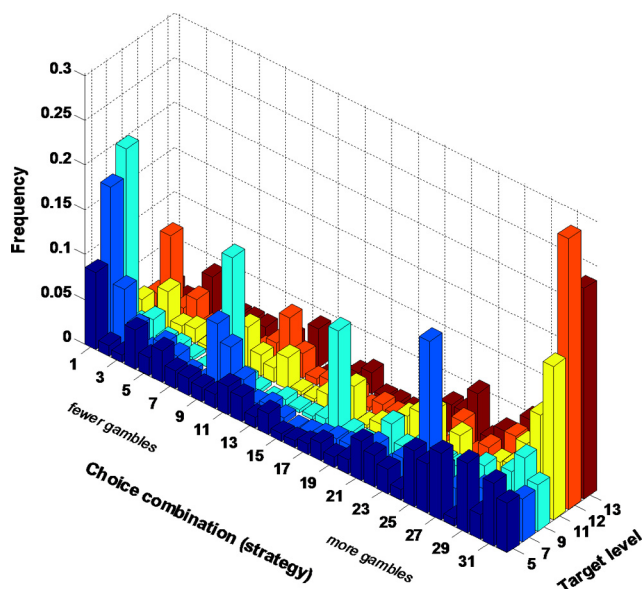
Figure 3. Choice strategies for different target levels, 16 subjects. Bar plot of frequencies with which different choice combinations (i.e., strategies) were chosen, for each of the six target levels used. Choice strategy on *x*-axis, numbered from 1 to 32 according to the $2^5 = 32$ possible combinations of five sequential binary choices to gamble or not gamble. Choice strategies sorted by increasing number of times a gamble is chosen (i.e., combination 1 in which all choices opt for sure amount, and combination 32 in which all choices opt for the gamble). Six target levels on *y*-axis. Choice frequency (i.e., number of blocks in which the specific choice combination occurred, divided by total number of blocks) on *z*-axis. Blockwise data are aggregated across all 16 subjects for illustration. Model fitting was performed on a subject-by-subject basis.

model obtains a mean criterion value of 78.3. The average distance (i.e., summed least-squares error) between the array of observed frequencies and the array of model-simulated frequencies is 6.3 for ACV and 78.3 for the random model. Therefore, the summed least-squares error for ACV is 92% less than is the case for the random model. According to the ACV model, all 16 subjects were averse to variance (variance coefficient ± SD, 0.21 ± 0.03) and were positive skew seeking (skewness coefficient ± SD, $1.2 \times 10^{-3} \pm 0.6 \times 10^{-3}$). The value of the σ parameter (temperature parameter of the softmax/logistic function used to account for noisy choice) were low (average ± SD σ, 0.31 ± 0.62). This indicates that the valuation model performs well at explaining choice without modeling a large degree of additional randomness in action selection. It is important to note that, although on average the ACV model was superior, there was heterogeneity in the best-fitting model on a subject-by-subject basis (Fig. 4*B*). The ACV model was superior to the SCV model in 13 of 16 subjects (Fig. 4*C*). Both parametric and nonparametric tests of the criterion value statistics at the group level revealed that ACV obtained a significantly better fit than SCV (paired *t* test, *p* = 0.001; binomial test, *p* = 0.002) but was indistinguishable from OCV on these behavioral data alone (paired *t* test, *p* = 0.294; binomial test, *p* = 0.402).

Functional imaging

We analyzed fMRI data, initially using the average continuation value model. A linear utility model is akin to the general linear model used in fMRI analysis, enabling us to decompose neural activity according to the effect of three statistical moments of the outcome distribution (mean, variance, and skewness). For the design matrix, we parametrically modulated the magnitude of the neural response on every trial with four regressors indicating

target level (high or low), expected value, variance, and skewness of the outcome distribution, respectively. The use of parametric modulators to model the neural response to complex stimuli with several dimensions is well established (Büchel et al., 1998; Wood et al., 2008), and the correlation of neural data with dynamically changing internal variables of a computational model has been implemented in several studies of the neural valuation system (O'Doherty et al., 2004; Samejima et al., 2005) (for review, see Corrado and Doya, 2007). We first include the target level to account for evoked activity differences solely attributable to changes in effort or concentration evoked by a difference between a low and high target and activity attributable to explicit tracking of the target or context. In addition, this removes correlations between the regressors induced by the fact that, at high targets, the expected value is naturally always low and the skewness is always positive (high chance of failing to reach target and receiving nothing).

We then performed a directed stepwise linear regression to analyze the contribution of each variable in turn to the BOLD signal, by sequentially orthogonalizing regressors. Because we had strong a priori regions of interest, derived from previous studies of single-shot decision-making for expected value and to a lesser extent variance and no a priori knowledge about skewness, we orthogonalized the regressors in this specific order. Thus, we first account for as much neural activity as possible with the expected value regressor, then explain residual activity with the variance regressor, and finally explain activity with the skewness regressor. Any residual activity correlating with skewness is therefore independent of expected value and variance. In addition, orthogonalization is necessary because correlations remain between the statistical moments even having accounted for gross differences attributable to the target level (correlation coefficients: expected value vs variance, 0.57; expected value vs skewness, 0.50; variance vs skewness, 0.01). It is important to note that including the target regressor changes our inference about activity tracking the predicted outcome statistics: we analyze activity tracking the conditional moments (i.e., expected value, variance, and skewness changes with respect to the current target level) rather than the raw unconditional statistics. This analysis conditional on current target is similar to previous studies investigating the tracking of value in different frames or conditions (De Martino et al., 2006, 2009; Elliott et al., 2008; Plassmann et al., 2008) and asks whether expected outcomes are encoded relative to context.

A key idea in this neural analysis is the principle of using neurophysiological data to arbitrate between models of decision-making that are difficult to distinguish using choice data alone. If future trials were not considered at all (i.e., participants were oblivious to the task structure and the need to attain a target) and instead if each lottery is compared with the sure amount in isolation, we would not expect to observe neural signals correlating with expected value or skewness (because all gambles were symmetric and had the same expected value) i.e., the null hypothesis. This hypothesis in itself is fairly trivial to refute using just behavioral data, because clearly our participants' choices are sensitive to the target level. Conversely, if individuals anticipate future outcomes in a strategic manner (using either of the two most likely strategies according to choice data, of OCV in which a specific set of choices are weighted, or ACV in which all possible future choices are weighted), the presence of such correlated signals can be interpreted as providing evidence that such strategies are taken into account and that the observed pattern of neural response tracks the anticipated distribution of outcomes (specif-

ically in brain regions previously implicated in representing statistical moments of choice in single choice paradigms). We aim to discriminate specifically between OCV and ACV models using neural data, given that these models were indistinguishable solely from our behavioral analysis of sequential choice. We use the fact that these models predict different anticipated outcome distributions on a trial-by-trial basis, together with fine-grained neural signal changes (as opposed to categorical choice data), to give additional power to arbitrate between models.

Average continuation value model
Fluctuations in expected value for each choice correlated with activity in right medial OFC (mOFC) (MNI coordinates: 6, 50, −14; $t = 4.16$, $p = 0.032$, small-volume corrected for region of interest) and nucleus accumbens (MNI coordinates: right nucleus accumbens, 4, 10, −6; $t = 3.74$, $p = 0.036$, small-volume corrected for regions of interest). In other words, activity in these regions tracked the orthogonalized component of expected value (meaning the error in a projection of expected value onto target level), according to our model of online tracking of outcome distributions. Note that this regressor is linearly independent of that tracking the target level (supplemental Table S1, available at www.jneurosci.org as supplemental material) (Fig. 5B). The target regressor itself correlated with activity in areas including right middle frontal gyrus (MNI coordinates: 46, −2, 54; $t = 5.76$, $p < 0.001$ uncorrected), anterior cingulate cortex (MNI coordinates: 6, 44, 14; $t = 5.08$, $p < 0.001$ uncorrected), and paracentral lobule/supplementary motor area (MNI coordinates; −6, −24, 56; $t = 4.66$, $p < 0.001$ uncorrected) (supplemental Table S2, available at www.jneurosci.org as supplemental material) (Fig. 5A). We also tested the alternative ACV model that did not explicitly model the target separately (i.e., we ask whether there is BOLD signal that correlates with the unconditional statistics of the outcome distribution, not adapted to target level). In this model, no brain activity positively correlated with the overall expected value of a choice, even at a liberal threshold of $p < 0.005$ uncorrected significance.

We next examined neural activity accountable by changes in the average variance of possible future outcomes given each choice, having accounted for activity attributable to target and expected value. The orthogonalized component of variance-related activity correlated with BOLD in anterior insula [MNI coordinates: right anterior insula, 40, 20, −6; $t = 3.64$, $p = 0.028$ small-volume corrected for regions of interest (supplemental data, available at www.jneurosci.org as supplemental material); left anterior insula, −38, 20, 4; $t = 3.84$, $p = 0.001$ uncorrected]), right putamen (MNI coordinates: 26, 28, −8; $t = 6.42$, $p < 0.001$ uncorrected), and right anterior cingulate cortex (MNI coordinates: 8, 44, 16; $p < 0.001$ uncorrected)
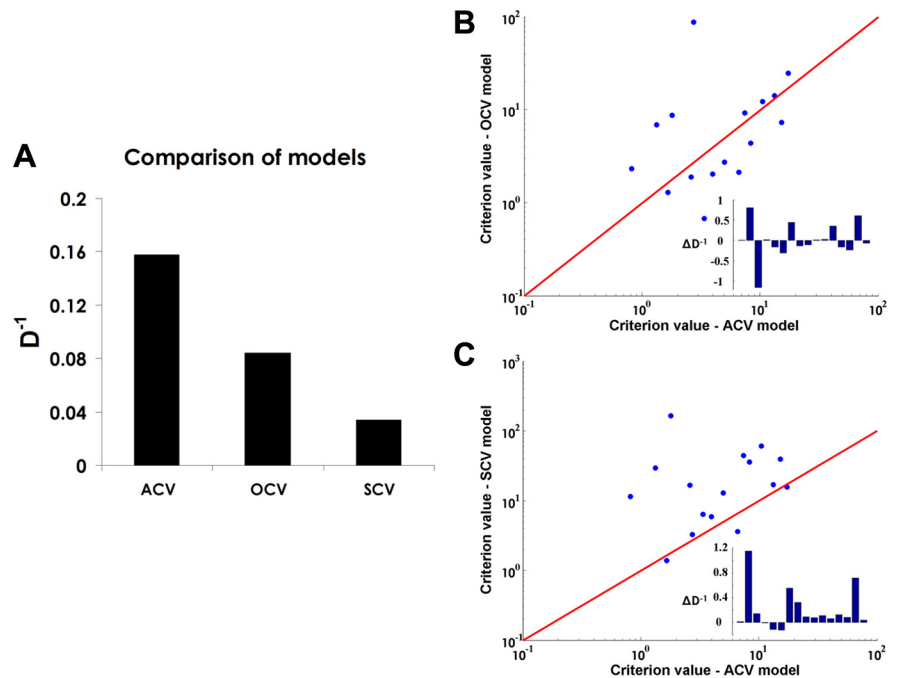

Figure 4. ***A***, Comparison of models I. There were three types of continuation value model, indicating different decision-making processes: optimal, in which only the relative optimal outcomes (of gambling or not gambling on each trial) are evaluated; average, in which the relative average of possible outcomes are evaluated (so all possible choice trajectories considered); and sure, in which the choice to gamble is evaluated with respect to the alternative of taking the sure option for the remainder of the block. To illustrate the relative superiority of the different models, we plot the inverse of the average criterion function [(mean *D*) $^{-1}$, where *D* is the criterion value (the distance between simulated and actual choice frequencies)]. Larger values for D^{-1} indicate a better model fit. ***B***, ***C***, Comparison of models II. For each subject ($n = 16$), the criterion value, *D*, from the mean–variance–skewness, average continuation value model, is compared with the optimal continuation value and sure continuation value models. Points above the red line indicate that the mean variance skewness model with average continuation value provides the better fit. Values are plotted on logarithmic axis for illustration; the red line indicates equality between models. The ACV model and OCV model are comparable in their behavioral fits across subjects. The ACV model outperforms the SCV model for most subjects (***C***). We also plot the residual values (differences) in D^{-1} per subject in the bottom corner.

(supplemental Table S3, available at www.jneurosci.org as supplemental material) (Fig. 5C).

Having accounted for neural activity attributable to target, expected value, and variance of anticipated outcomes, we next sought to explain residual activity in terms of the (orthogonalized component of) skewness of the expected outcome distribution (calculated as the expected cubed deviations from expected outcomes). We observed activity correlating with skewness in medial frontal pole, left superior parietal cortex and postcentral gyrus, and left inferior frontal gyrus ($p < 0.001$ uncorrected) (supplemental Table S4, available at www.jneurosci.org as supplemental material) (Fig. 5D).

As an additional analysis, we estimated a separate general linear model in which we modeled neural responses covarying with subject-specific expected utility on a trial-by-trial basis, calculated according to the behavioral parameters estimated from subjects' parameters from the ACV model. As might be expected, the largest cluster of significant activity correlating with expected utility was found in medial prefrontal cortex (peak voxel MNI coordinates: −8, 56, −2; $t = 3.87$; extent = 66 voxels) (supplemental Table S5, available at www.jneurosci.org as supplemental material).

Optimal continuation value model
Given that the OCV model was statistically indistinguishable from ACV on the behavioral data alone, we implemented an analysis based on predictions from the alternative OCV model for
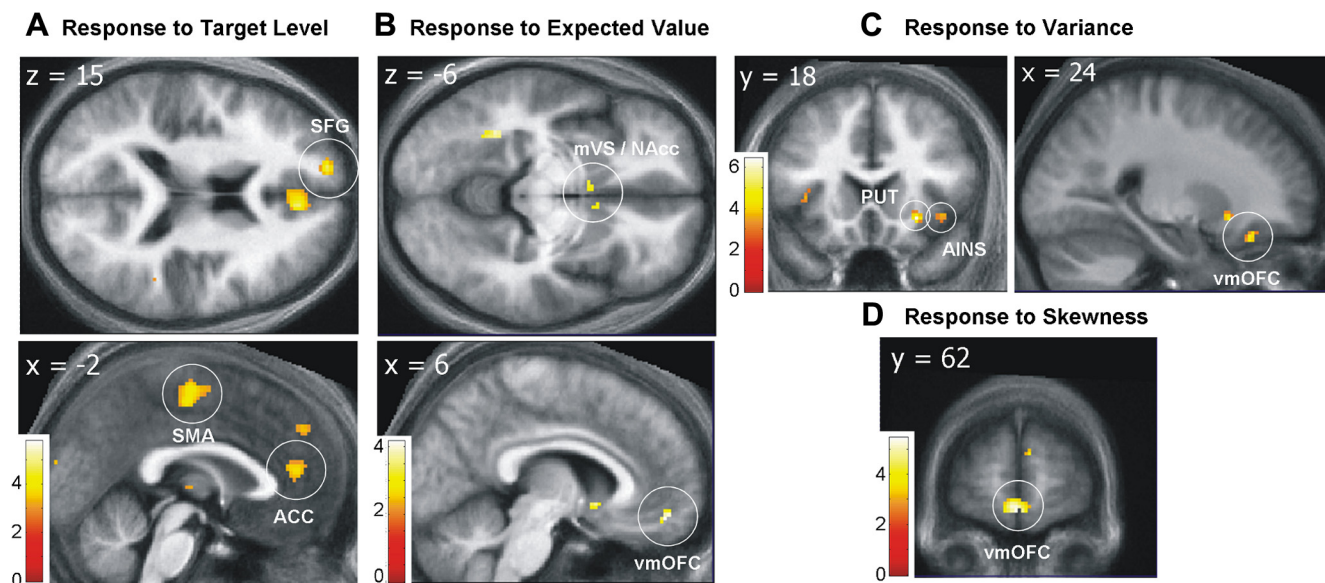
Figure 5. Responses to target level, expected value, variance, and skewness. **A**, Main effect of target level. SMA, Supplementary motor area; ACC, anterior cingulated cortex; SFG, superior frontal gyrus. **B**, Main effect of parametric response to expected value given by ACV model. mVS/NAcc, Medial ventral striatum/nucleus accumbens; VmOFC, ventromedial orbitofrontal cortex. **C**, Parametric response to expected variance (risk) of each choice. PUT, Putamen; AINS, anterior insula. **D**, Parametric response increasing with (positive) skewness in medial frontal pole. SPM *t* thresholded at $p < 0.005$ superimposed on a canonical structural template. Color bars show *t* value scales for voxel color maps.

each subject, to investigate whether neural activity correlated with the internal parameters of this model. In essence, we are asking whether brain activity can adjudicate between models. We formulated the fMRI design matrix in an identical manner, modeling activity as a series of spike events, with their heights modulated by parametric regressors corresponding to target level, expected value, variance, and skewness of the outcome distribution under a set optimal strategy but contingent on previous decisions in a block. As above, inferences were made at the group level. There was no significant activity in a priori regions of interest correlating with the regressors tracking the outcome distribution, at a threshold of $p < 0.001$ uncorrected significance. For completeness, we tested the alternative OCV model not explicitly modeling the target separately and also found that no brain activity positively correlated with the overall expected value of a choice (at $p < 0.005$ uncorrected). This suggests that neural activity in brain regions previously associated with economic decision-making is better captured by an ACV model, with on-line trial-by-trial updating rather than the OCV model with a fixed pre-set strategy.

Discussion

We first asked how humans evaluate outcome distributions from different strategies in a sequential choice task. Using a mean–variance–skewness model, we find that our subjects are variance averse and positive-skew seeking. Positive-skew-seeking manifests when participants excessively opt for the sure rather than risky option even with low targets (in which the chance of failing to reach the target is small). This implies an attraction of small chances of above-average outcomes and dislike of small-probability below-average outcomes. In effect, we observe a preference for relative gains over losses, similar to prospect theory.

fMRI data revealed brain activity correlating with the statistical moments of a distribution of outcomes in prototypical valuation and risk-sensitive areas. Previous studies of risky decisions have segregated risk and value-related activity in regions such as cingulate and insula cortices (risk) and ventral striatal and medial

orbitofrontal areas (valuation) (Kuhnen and Knutson, 2005; Lee, 2005; Huettel et al., 2006; Rangel et al., 2008). Finding separable areas of brain activity parametrically varying with mean, variance, and skewness supports predictions from an MVS model, but we cannot rule out alternative neural implementations of subjective utility. It is possible that the variance-related activity we see may reflect downstream emotional or physiological responses consequential on detecting increased risk rather than direct risk-assessment itself. For example, the insula supports introceptive processing, and activity we observe could be explained either as risk representation or as an arousal response consequent on the perception of risk (Critchley et al., 2004). However, a parametric response to risk in these regions is less easily explained as an arousal response unless one invokes a monotonic relationship between risk and arousal.

Our finding that expected value correlates with activity in mOFC supports an hypothesis that mOFC integrates overall value given a predicted distribution of outcomes. In our task these variables related to a distribution of outcomes not from a single choice but from a set of serial choices, forecast to the end of each block. This corroborates a suggestion that neural response to value invokes an integrated, goal-directed, representation of choice (Quintana and Fuster, 1999; Fincham et al., 2002).

Our analysis includes a regressor accounting for the target (high/low), controlling for target-induced changes in attention, concentration, or effort, and means that OFC responses track the moments of the outcome distribution conditional on target level. In effect, the mOFC BOLD signal tracks relative rather than absolute changes in expected value (because we did not see similar activity without controlling for target level). This suggests adaptive value encoding, similar to findings from direct neuronal recordings in monkeys in which a proportion of OFC neurons adapt to condition, manifesting similar ranges of response under different scales of outcomes (Padoa-Schioppa and Assad, 2008; Padoa-Schioppa, 2009; Kobayashi et al., 2010). This adaptation potentially can overcome the limited dynamic range of neuronal

signaling, implying that responses to expected value are integrated with information about the target in generating action values.

We also identify a skewness response, comprising medial prefrontal and superior parietal cortex, also shown to reflect subjective value in tasks with stochastic outcomes (Peters and Büchel, 2009). It is unlikely that this simply reflects cognitive demand or planning, because we see a parametric response even having separately accounted for the target level. We detect a signal reflecting expected utility of each choice (incorporating the target, relative expected value, and risk), calculated as the subject-specific combination of these statistical moments, in medial prefrontal cortex. This locus of activity overlaps with areas in which activity correlates with subjective utility (Daw et al., 2006). However, computations for sequential decision-making, in which outcomes are stochastic and forecast several trials into the future, are represented in a more anterior location to that found when decision utilities represent deterministic outcomes from single-shot choices (Plassmann et al., 2007).

Our second question related to how individuals account for their possible future choices when selecting actions. Behaviorally, we found individuals reevaluate on each trial (ACV) rather than comparing choices with the risk-free alternative (SCV). Furthermore, neural activity distinguished between behaviorally equivalent OCV and ACV models, with correlations of key variables from the latter rather than the former. However, classical dynamic programming models are based on OCV, in which decision-makers assume a specific optimal series of future planned choices. These models insist on dynamic consistency (choices are independent of the order in which options are presented) and have been used to describe choice in computational (Sutton and Barto, 1998), ecological (Houston et al., 1988), and economic (Samuelson, 1969) settings. ACV implies that potential outcomes from a number of strategies influence current choice in the decision-making process (because ACV decision-makers assume future choices are made randomly).

The likelihood that decision-makers represent or weight outcomes of alternative strategies relates to the possibility that future actions may deviate from an optimal trajectory. This can be either intentional (as a result of exploration or future constraints on available choices) or by accident (lapses or mistakes). In reality, we are unlikely to follow a predetermined path in our strategic decisions. If we ignored alternative outcomes altogether, then deviations from an optimal strategy would lead to unpredicted and possibly far worse outcomes than originally envisaged. Indeed, there is good evidence that weighting of even potentially irrelevant alternative outcomes plays a role in paradoxes of choice (Allais, 1953; Loomes and Sugden, 1982; Birnbaum, 2008), with counterfactual outcomes being represented in prefrontal cortex (Ursu and Carter, 2005) and striatum (Lohrenz et al., 2007). An alternative reason why an "average plan" rather than an optimal strategy might be used is because of additional mental effort required in planning future actions. Predicted outcomes could instead be sampled from the whole range of possible alternatives to build up an average picture of what might transpire given current choice. ACV far outperforms random choice, which would be an alternative best heuristic if decision-makers were completely ignorant of future options. Instead, ACV explicitly models the assumption that all future options are known, but that despite this the decision-maker does not have a specific plan of their future choices.

Although using a fixed strategy model is possible in our task, it is likely to dramatically fail in situations in which an individual errs or if some planned alternatives are no longer available. An ACV decision-maker considering all possible future outcomes is myopic (i.e., does not deterministically plan choices in advance). However, such a decision-maker can mitigate future errors or constraints by weighting all possible action–outcome combinations, enabling recovery from error by selecting the best set of remaining choices without needing to assume a fixed strategy. In other words, it makes sense that dynamic programming should account for a decision-maker's awareness that deviations from a specific policy may occur in the future. One method of implementing this is to optimize the average continuation value. ACV can also partly capture decision processes in which a proportion of (but not necessarily all) possible outcomes in a decision tree are considered. A more informed version of the ACV model might weight strategies according to the proportion of time that they are expected to be chosen either according to previous experience or based on a rational expectations model similar to quantal response equilibria models of choice (McKelvey and Palfrey, 1998). We necessarily test the joint hypothesis (that both MVS and ACV models are true), because our continuation value models are coupled to a valuation model. Thus, although the neural data support ACV over OCV contingent on MVS, it is possible that this could be bettered by a combination with an alternative utility model. We cannot draw direct statistical comparisons between utility models in the current framework (because the models are non-nested with different numbers of parameters), and these additional model variants remain to be tested in future work. It is possible that we do not see neural responses corresponding to an OCV model even in individuals who actually do follow this strategy purely because there is no requirement for a trial-by-trial tracking of the outcome distribution if you have preplanned a trajectory of choices. However, the fact that we see responses corresponding to the ACV model suggests that, at least in some subjects, these average continuation values are being continuously tracked and reevaluated.

In conclusion, we provide behavioral and neural data showing how humans make sequential decisions, a central element in decision-making scenarios ranging from foraging to financial investment. Our design assumes a fixed best-fitting strategy across subjects and cannot rule out variation in strategies both between and within subjects (i.e., switching strategies through the experimental session). However, such heterogeneity would have the effect of obscuring our ability to differentiate between models. Rather than conforming to standard models of sequential decision-making, our data suggest that a set of possible strategies are neurally represented and drive choices. More generally, our findings indicate that strategic outcomes are evaluated by similar neural metrics as in single-shot choice, in which a behavioral preference for higher-order features of outcome distributions is mirrored by neural sensitivity to expected value, variance, and skewness. Thus, it seems that phylogenetically ancient circuitry subserving valuation and reward also enables the sophisticated representation of the future and its alternatives.

References

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. Neuroimage 31:790–795.

Allais M (1953) Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. Econometrica 21:503–546.

Birnbaum MH (2008) New paradoxes of risky decision making. Psychol Rev 115:463–501.

Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus-

response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8:140–148.

Corrado G, Doya K (2007) Understanding neural coding through the model-based analysis of decision making. J Neurosci 27:8178–8180.

Critchley HD, Wiens S, Rotstein P, Ohman A, Dolan RJ (2004) Neural systems supporting interoceptive awareness. Nat Neurosci 7:189–195.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441:876–879.

Dekel E, Lipman BL, Rustichini A (1998) Standard state-space models preclude unawareness. Econometrica 66:159–173.

De Martino B, Kumaran D, Seymour B, Dolan RJ (2006) Frames, biases, and rational decision-making in the human brain. Science 313:684–687.

De Martino B, Kumaran D, Holt B, Dolan RJ (2009) The neurobiology of reference-dependent value computation. J Neurosci 29:3833–3842.

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.

Epstein LG, Schneider M (2003) Recursive multiple-priors. J Econ Theory 113:1–31.

Fincham JM, Carter CS, van Veen V, Stenger VA, Anderson JR (2002) Neural mechanisms of planning: a computational analysis using event-related fMRI. Proc Natl Acad Sci U S A 99:3346–3351.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. Neuroimage 7:30–40.

Hansen LP (1982) Large sample properties of generalized method of moments estimators. Econometrica 50:1029–1054.

Houston A, Clark C, McNamara J, Mangel M (1988) Dynamic models in behavioural and evolutionary ecology. Nature 332:29–34.

Huettel SA, Stowe CJ, Gordon EM, Warner BT, Platt ML (2006) Neural signatures of economic preferences for risk and ambiguity. Neuron 49:765–775.

Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R (2002) Image distortion correction in fMRI: a quantitative evaluation. Neuroimage 16:217–240.

Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263–291.

Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. J Neurosci 25:4806–4812.

Kobayashi S, Pinto de Carvalho O, Schultz W (2010) Adaptation of reward sensitivity in orbitofrontal neurons. J Neurosci 30:534–544.

Kuhnen CM, Knutson B (2005) The neural basis of financial risk taking. Neuron 47:763–770.

Lee D (2005) Neuroeconomics: making risky choices in the brain. Nat Neurosci 8:1129–1130.

Lohrenz T, McCabe K, Camerer CF, Montague PR (2007) Neural signature of fictive learning signals in a sequential investment task. Proc Natl Acad Sci U S A 104:9493–9498.

Loomes G, Sugden R (1982) Regret theory: an alternative theory of rational choice under uncertainty. Econom J 92:805–824.

Machina MJ (1989) Dynamic consistency and non-expected utility models of choice under uncertainty. J Econ Lit 27:1622–1668.

Markowitz H (1952) Portfolio selection. J Finance 7:77–91.

Mazziotta J (2001) A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc B Biol Sci 356:1293–1322.

McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica 57:995–1026.

McKelvey RD, Palfrey TR (1998) Quantal response equilibria for extensive form games. Exp Economics 1:9–41.

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.

Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. J Neurosci 29:14004–14014.

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.

Panageas S, Westerfield MM (2009) High-water marks: high risk appetites? Convex compensation, long horizons, and portfolio choice. J Finance 64:1–36.

Peters J, Büchel C (2009) Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. J Neurosci 29:15727–15734.

Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. J Neurosci 27:9984–9988.

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.

Quintana J, Fuster JM (1999) From perception to action: temporal integrative functions of prefrontal and parietal neurons. Cereb Cortex 9:213–221.

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

Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310:1337–1340.

Samuelson PA (1969) Lifetime portfolio selection by dynamic stochastic programming. Rev Econ Stat 51:239–246.

Stephens DW, Charnov EL (1982) Optimal foraging: some simple stochastic models. Behav Ecol Sociobiol 10:251–263.

Stephens DW, Brown JS, Ydenberg RC (2007) Foraging: behavior and ecology. Chicago: University of Chicago.

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: Massachusetts Institute of Technology.

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.

Weiskopf N, Hutton C, Josephs O, Turner R, Deichmann R (2007) Optimized EPI for fMRI studies of the orbitofrontal cortex: compensation of susceptibility-induced gradients in the readout direction. Magn Reson Mater Phys Biol Med 20:39–49.

Wood G, Nuerk HC, Sturm D, Willmes K (2008) Using parametric regressors to disentangle properties of multi-feature processes. Behav Brain Funct 4:38.

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.