

Deciding When to Decide: Time-Variant Sequential Sampling Models Explain the Emergence of Value-Based Decisions in the Human Brain

Sebastian Gluth,¹ Jörg Rieskamp,² and Christian Büchel¹

¹Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany and ²Department of Psychology, University of Basel, 4055 Basel, Switzerland

The cognitive and neuronal mechanisms of perceptual decision making have been successfully linked to sequential sampling models. These models describe the decision process as a gradual accumulation of sensory evidence over time. The temporal evolution of economic choices, however, remains largely unexplored. We tested whether sequential sampling models help to understand the formation of value-based decisions in terms of behavior and brain responses. We used functional magnetic resonance imaging (fMRI) to measure brain activity while human participants performed a buying task in which they freely decided upon how and when to choose. Behavior was accurately predicted by a time-variant sequential sampling model that uses a decreasing rather than fixed decision threshold to estimate the time point of the decision. Presupplementary motor area, caudate nucleus, and anterior insula activation was associated with the accumulation of evidence over time. Furthermore, at the beginning of the decision process the fMRI signal in these regions accounted for trial-by-trial deviations from behavioral model predictions: relatively high activation preceded relatively early responses. The updating of value information was correlated with signals in the ventromedial prefrontal cortex, left and right orbitofrontal cortex, and ventral striatum but also in the primary motor cortex well before the response itself. Our results support a view of value-based decisions as emerging from sequential sampling of evidence and suggest a close link between the accumulation process and activity in the motor system when people are free to respond at any time.

Introduction

A value-based decision is a deliberative process that requires the ability and the time to evaluate the attractiveness of a particular choice option (Gold and Shadlen, 2007; Rangel et al., 2008). Consider an everyday situation such as booking a hotel room via the Internet. The hotel you are evaluating at the moment is very appealing, the location good, and the price low. You are already tempted to go for it, but then you decide to also check the reviews from previous customers. You read that the personnel is unfriendly, the breakfast poor, and the bathrooms unbearable. Many reviewers do not recommend the hotel. Your initial conviction has disappeared.

The described decision process can be characterized as a time-consuming evidence accumulation process, which constitutes the framework for sequential sampling models (SSMs) of decision making (Ratcliff, 1978; Busemeyer and Townsend, 1993; Usher

and McClelland, 2001). Information about choice options is repeatedly sampled over time and accumulated into a preference state. If this preference state exceeds a decision threshold, a decision is made (e.g., to book the hotel). SSMs have advanced our knowledge on the neuronal mechanism of perceptual decisions in humans and nonhuman primates (Gold and Shadlen, 2007; Heekeren et al., 2008). Neurons in the lateral intraparietal area of the monkey brain apparently encode the decision variable (DV), as it evolves during the decision process (Platt and Glimcher, 1999; Roitman and Shadlen, 2002). Neuroimaging studies suggest a corticostriatal circuit to mediate the speed–accuracy trade-off in perceptual decisions by adjusting the decision threshold at the onset of the accumulation process (Bogacz et al., 2010). The temporal characteristics of value-based decisions are, however, less understood.

The purpose of the present study was to employ SSMs for describing the emergence of value-based decisions over time and to elucidate the neuronal basis underlying the cognitive mechanism. We thereto designed a task (Fig. 1A), in which participants could either buy or reject stock offers of unknown value. Rating companies provided probabilistic information about the stock's value, but sampling this information was coupled with a fixed cost. Contrary to previous sequential decision-making tasks (Yang and Shadlen, 2007; de Lange et al., 2010; Stern et al., 2010), participants were free to respond at any time. This design allowed us to conceptualize and test computational models against each other in predicting how and when

Received Feb. 15, 2012; revised June 1, 2012; accepted June 7, 2012.

Author contributions: S.G., J.R., and C.B. designed research; S.G. performed research; S.G., J.R., and C.B. analyzed data; S.G., J.R., and C.B. wrote the paper.

This research was supported by grants from Deutsche Forschungsgemeinschaft (GRK 1247, CINACS) and Bundesministerium für Bildung und Forschung (01GQ0912, Bernstein Focus Learning). J.R. was supported by a grant from the Swiss National Science Foundation (SNSF 100014_130149/1).

The authors declare no competing financial interests.

Correspondence should be addressed to Sebastian Gluth, Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. E-mail: s.gluth@uke.de.

DOI:10.1523/JNEUROSCI.0727-12.2012

Copyright © 2012 the authors 0270-6474/12/3210686-13\$15.00/0

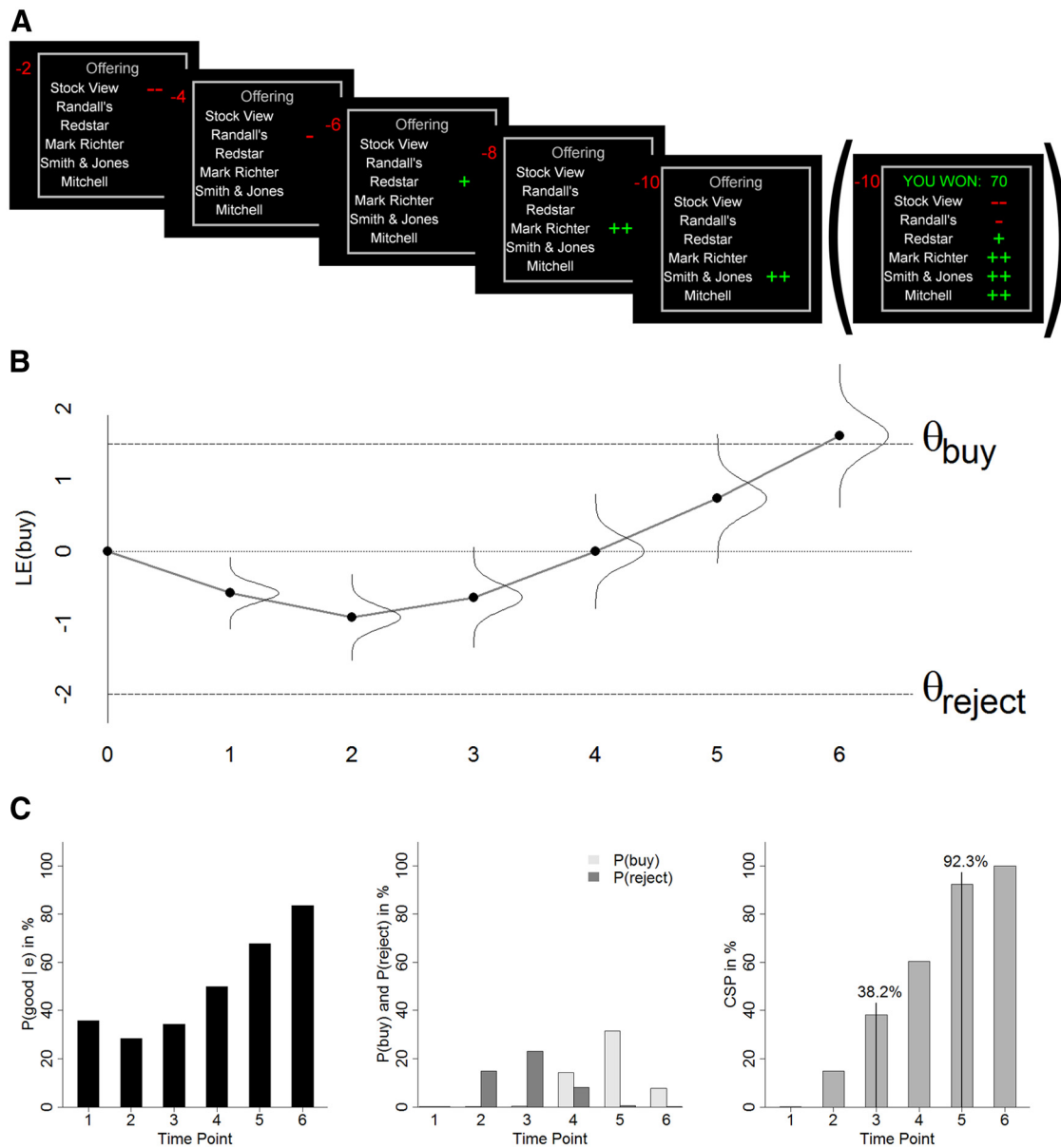


Figure 1. Experimental design and computational modeling. **A**, In each trial, participants could buy or reject a stock of unknown value. Six rating companies provided probabilistic information. Ratings followed each other sequentially with presentation time varying between 2 and 4 s. Participants could respond at any time (decision point in the example is $t = 5$) and lost points for awaiting further ratings (upper left corner). The feedback slide (in brackets) was only shown during training. **B**, Illustration of SSMs for the current task design. The log-evidence for buying the stock is updated at each rating presentation and used as decision variable for the random walk, which starts at 0. Two boundaries for the choice options (buy, reject) delimit the process. Note that the distance between starting point and boundary can differ between θ_{buy} and θ_{reject} . The time-variant SSM assumes decreasing rather than fixed boundaries (see Materials and Methods). **C**, The sequence of ratings determines how the probability of a good offer evolves over time (left). In the context of model fitting, subject- and trial-specific probabilities of buying and rejecting at each time point are obtained (middle). The CSP (right) gives the probability that a decision has been made until a specific time point and is used to assess trial-by-trial variations in response tendencies (Fig. 7C). Note that the graphs shown in **C** refer to the specific combination of ratings shown in **A**. Response probabilities and CSP vary from trial to trial and between participants.

participants made their choices. Model parameters of the best model were then used to inform our functional magnetic resonance imaging (fMRI) analyses. We hypothesized that ventral striatum (VS) and ventromedial prefrontal cortex (vmPFC) track the updated value information (O’Doherty et al., 2004; Kable and Glimcher, 2007; Wallis, 2007). This information should be immediately accessible for motor preparation and output regions such that the gradual buildup of DVs will be reflected in the motor system (Donner et al., 2009; Cisek and Kalaska, 2010). Furthermore, based on findings from perceptual decision making, we hypothesized that trial-by-trial fluctuations in the decision threshold are related to fMRI signal variations in the caudate

nucleus and the presupplementary motor area (pre-SMA; Bogacz et al., 2010; van Maanen et al., 2011).

Materials and Methods

Participants. Participants were 29 right-handed healthy persons with normal or corrected-to-normal vision. Three participants were excluded from the analysis, because for two participants fMRI data could not be analyzed due to severe MR artifacts and one participant did not work on the task with sufficient commitment (the participant reported that she tried to finish the study as quickly as possible). Thus, the final sample included 26 participants (mean age = 25.4 years, ± 3.6 SD, 21–36 years; 12 females). The study was approved by the local ethics committee and all

participants gave written informed consent. Participants were reimbursed for participation and could earn additional money by winning points in the task.

Experimental design. In each trial, participants were offered a stock and had to decide whether to buy or reject the offer (Fig. 1A). A gray frame enclosing the heading “offering” and the names of six fictitious rating companies were presented on the screen throughout the entire fMRI experiment. In addition, a counter at the upper left side of the frame depicted the costs for awaiting rating information and was set to a white-colored “0” between trials. A trial commenced with the rating of the first company appearing next to the first company’s name. Positive ratings were colored green; negative ratings were colored red. At the same time, the counter turned to a red-colored “−2” (each rating cost 2 points). After a variable delay (2, 2.5, 3, 3.5, or 4 s) the second rating was displayed next to the second company’s name, the counter turned to “−4,” and the previous rating disappeared. This procedure continued until the last rating was presented (again for at least 2 s) or a response was made, which terminated the trial (i.e., the current rating disappeared and the counter was set back to “0”; feedback was only provided during prescan training). Trials were separated by a variable interval of 2–9 s. The order of rating companies was fixed from top to bottom.

Participants were told that stocks were either good (value: +80 points) or bad (value: −80 points) and buying a stock would lead to the payment of its value. Participants were instructed to respond whenever they wanted during the trial, but that a response had to be given after disclosure of the sixth rating at the latest (otherwise they would receive −92 points). They were further informed about the possible ratings (“−,” “+,” “+,” or “+ +”), the costs for each rating, the independence of ratings of different companies from each other, that all companies were equally important, and that the ratings contained probabilistic information about stock values. They were instructed that the probability of being offered a good stock is increased to 60% given a “+” and decreased to 40% given a “−,” that “+ +” (“− −”) ratings are equivalent to two separate “+” (“−”) ratings, and that in general the more +’s and the less −’s presented, the higher the probability is of a good stock. In fact, the probability of a good stock given the entire accumulated evidence e at time point t is updated with every new rating according to Bayes’ theorem (compare Busemeyer and Pleskac, 2009) as in the following:

$$P(\text{good} | e)_t = \frac{P(\text{good} | e)_{t-1} \times P(\text{rating} | \text{good})_t}{P(\text{good} | e)_{t-1} \times P(\text{rating} | \text{good})_t + P(\text{bad} | e)_{t-1} \times P(\text{rating} | \text{good})_t} \quad (1)$$

where $P(\text{good} | e)_{t-1}$ is the prior probability of a good stock given the previous evidence [note that $P(\text{bad} | e)_{t-1} = 1 - P(\text{good} | e)_{t-1}$] and $P(\text{rating} | \text{good})_t$ is the likelihood of the current rating (“−,” “+,” “+,” or “+ +”) given that the stock is good. Note that $P(e | \text{good})_t = P(\text{good} | e)_t$, when all six cues are known (i.e., at $t = 6$) (Yang and Shadlen, 2007; Philiastrides et al., 2010). Since the prior probability for good and bad stocks is equal and since $P(\text{good} | “+”) = 1 - P(\text{good} | “−”) = 1 - P(\text{bad} | “+”) = P(\text{bad} | “−”) = 60\%$, Equation 1 can be simplified to:

$$P(\text{good} | e)_t = \frac{P(\text{good} | “+”)^{S_t}}{P(\text{good} | “+”)^{S_t} + [1 - P(\text{good} | “+”)^{S_t}]} = \frac{0.6^{S_t}}{0.6^{S_t} + 0.4^{S_t}} \quad (2)$$

with

$$S_t = \sum_{i=1}^t \text{rating}_i \quad (3)$$

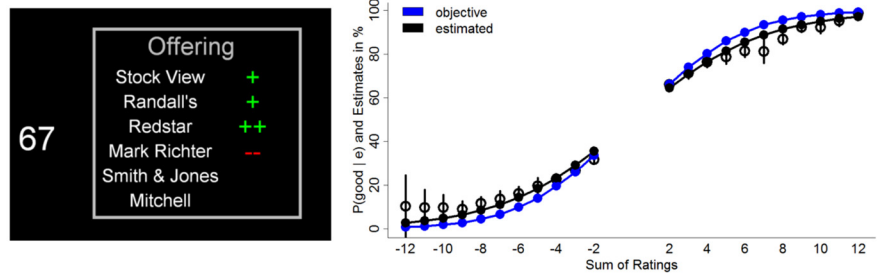


Figure 2. Probability estimation task and results. After scanning, participants estimated the probability of a good stock $P(\text{good} | e)$, given a selected combination of ratings (left). Results (right) show that participants are able to approximate $P(\text{good} | e)_t$ with a tendency to overestimate low and underestimate high probabilities. The black line is the average logistic regression curve fitted to the probability estimates (black circles with SEM bars) for each participant, which are close to the objective values (blue line with dots).

where rating_i is the rating of the company i (i.e., $\text{rating} = -2$ for “− −,” $= -1$ for “−,” $= 1$ for “+,” and $= 2$ for “+ +”); therefore S_t is the sum of all ratings presented until t . Equations 2 and 3 can be used to calculate the posterior probability when all cues are known. Calculating partial evidence given only 1–5 cues is more complex due to conditional dependencies (Yang and Shadlen, 2007). We followed Yang and Shadlen (2007) in deriving partial $P(\text{good} | e)_t$ directly from the set of all possible rating combinations. Although we do not think that people can exactly determine $P(\text{good} | e)_t$ after each rating, we assume that people can at least approximate it. To confirm this assumption we performed a probability estimation judgment task, described below, after the scanning session in which the participants had to estimate $P(\text{good} | e)_t$ for exemplary rating configurations (Fig. 2). Further note that we tested various SSMs, described below, employing different DVs including a model that simply relies on the sum of ratings as specified in Equation 3 (see Fig. 4B).

Overall, 120 stocks were offered during the fMRI experiment. Rating configurations for these stocks were generated randomly for each participant with the restriction that each of the four possible ratings of the first company was presented in exactly 25% of the trials. The length of the experiment depended on the amount of awaited ratings but did not exceed 40 min. Participants knew that the number of stocks was fixed and that faster play would not lead to getting more offers. Two practice sessions of 32 trials each preceded the fMRI experiment: in the first session, feedback was given after each trial to familiarize participants with the probabilistic nature of rating information and in the second practice session and the scanning session no feedback was provided.

Probability estimation task and questionnaire. After scanning, participants worked on a second task that examined their ability to estimate $P(\text{good} | e)_t$. In every trial, participants saw a selected combination of ratings on a computer screen and were asked to indicate the probability that the current offer was *good* given the presented ratings by typing in their estimate in percentages (integers from 0 to 100) (Fig. 2). The number of presented ratings per combinations varied from 1 to 6 and the sum of all ratings varied from −12 to −2 and from 2 to 12. Within these ranges, combinations for all possible sums for all time points were presented, resulting in 72 trials. At $t = 1$, there are two possible sums (−2 and 2) and thus two trials; at $t = 2$, there are six possible sums (−4, −3, −2, 2, 3, and 4) and thus six trials, etc. There was no time limit for this task. At the end of the study, participants completed a computerized version of the Temperament Character Inventory (TCI) questionnaire (Cloninger, 1994) including only a selected number of items: 21 novelty-seeking items (11 of subscale exploratory excitability, 10 of subscale impulsiveness), 18 harm avoidance items (11 of subscale anticipatory worry, 7 of subscale fear of uncertainty), and 15 cooperativeness items (8 of subscale social acceptance, 7 of subscale empathy).

Computational models (I). In the following, we describe the computational models that were applied for predicting behavior. Except for the optimal solution, the different models can be expressed by means of the SSM approach and we therefore outline this approach first. As illustrated in Figure 1B, a DV evolves during each trial and is linked to the consecutive presentation of ratings. Importantly, we assume that the accumu-

lation process is affected by Gaussian noise and hence is probabilistic rather than deterministic. It follows that the DV is not a point estimate but specifies the mean of a normal distribution of possible values for this DV with variance $\sigma^2 \times t$ (Cox and Miller, 1965), where σ is a free parameter. A stock is bought as soon as the DV crosses an upper boundary or threshold θ_{buy} and rejected as soon as it crosses a lower boundary θ_{reject} , where θ_{buy} and θ_{reject} are two free parameters. An even simpler version could assume $\theta_{reject} = -\theta_{buy}$, reducing the number of free parameters by one. However, our modeling results indicated that all models perform better when using asymmetric rather than symmetric boundaries for the two choice options even after punishing for the additional free parameter. For parameter estimation we identified the model's predicted probability of the specific responses (buy vs reject) at any of the six rating presentations. This probability is determined by the probability that the DV crosses one of the boundaries at t :

$$P(buy | t) = P(DV_t > \theta_{buy}) = \Phi\left(\frac{-\theta_{buy} + DV_t}{\sigma \times \sqrt{t}}\right) \quad (4)$$

$$P(reject | t) = P(DV_t < \theta_{reject}) = \Phi\left(\frac{\theta_{reject} - DV_t}{\sigma \times \sqrt{t}}\right) \quad (5)$$

where $\Phi(x)$ refers to the standard normal cumulative density function at x . The last rating forms a special case, because participants knew that they were punished for not responding at all (in fact, only one participant failed to respond in only one trial). Therefore we normalized the probability of either buying or rejecting at $t = 6$, so that the normalized probabilities $P(buy | t = 6)$ and $P(reject | t = 6)$ add up to 1, while the ratio between the probabilities of buying and rejecting remained unaltered. Importantly, Equations 4 and 5 still do not define the probability of a response at t , as the probability that a response could have been made earlier than t is ignored. That is, $P(buy | t)$ is the conditional probability of buying at t given that a response had not been made earlier. Taking this conditionality into account, we arrive at:

$$P(choice)_t = P(choice | t) \times \prod_{i=1}^{t-1} [1 - P(buy | i) + P(reject | i)] \quad (6)$$

where *choice* refers to either buy or reject.

The general SSM approach described so far is equivalent to our first computational model, the standard SSM. To specify the DV for this model we followed previous work (Yang and Shadlen, 2007; Philastides et al., 2010) and calculated the log-likelihood ratio that the current offer is good (based on $P(good | e)_t$; see Equation 1), which we refer to as the log-evidence for buying the stock:

$$DV_t^{SSM} = LE(buy)_t = \ln \frac{P(good | e)_t}{1 - P(good | e)_t} \quad (7)$$

To test whether participants were indeed updating information we tested this model against a second model that used only the current log-evidence as DV, the current evidence model:

$$DV_t^{current\ evidence} = \ln \frac{P(rating | good)_t}{1 - P(rating | good)_t} \quad (8)$$

This model thus uses rating information for making predictions but disregards the sequential nature of the task. Alternatively, participants might integrate rating information over time but tend to forget about previous ratings. We tested such an SSM with forgetting by introducing an additional free parameter ω ($0 < \omega < 1$; see also Busemeyer and Townsend, 1993) when defining the DV:

$$DV_t^{SSM\ with\ forgetting} = \sum_{i=1}^t \ln \frac{P(rating | good)_i}{1 - P(rating | good)_i} \times (1 - \omega)^{t-i} \quad (9)$$

such that evidence is accumulated (as indicated by the Σ sign) but is weighted less, the more it dates back in time (i.e., the larger $t - i$). As another alternative to the SSM approach, one could assume that there is

not a single continuous decision process during a trial but that each presentation of a new rating induces a new decision process. Effectively, such a model is very similar to the standard SSM except that it would suggest a fixed rather than increasing variance in the estimation of the DV. We therefore denoted this model as fix variance and realized it by taking out \sqrt{t} from the denominator in Equations 4 and 5.

The four models described so far assume fixed decision boundaries throughout a trial. Previous work, however, suggested that boundaries might decrease over time to ensure that the random walk will eventually cross one of the boundaries and the decision process does not continue for an unreasonable amount of time (Ditterich, 2006a,b; Churchland et al., 2008; Cisek et al., 2009). Accordingly, we also tested a time-variant SSM with linearly decreasing boundaries ("decreasing" means that the distance to 0 is reduced):

$$\theta_{buy,t} = \theta_{buy} - \lambda \times t \quad (10)$$

$$\theta_{reject,t} = \theta_{reject} + \lambda \times t \quad (11)$$

where λ is an additional free parameter controlling the strength of decrease. Note that this parameter has also been interpreted as an increasing urgency signal (Cisek et al., 2009).

Finally, we tested a model that arrives at the optimal solution. To perform optimally in the task, one would have to realize that there are three choice options at each point of time (i.e., to buy, to reject, or to wait for further ratings) to calculate their respective expected values (EV; the information given to participants is sufficient to estimate all of them) and to strictly select the option with the highest EV (Busemeyer and Pleskac, 2009). Note that by coupling the disclosure of ratings with a fixed cost, we were able to set the average optimal decision point in the middle (instead of at the end) of the sequence of ratings. A stock should be bought as soon as its EV is higher than the EVs for rejecting and waiting and rejected as soon as the EV for rejecting (equivalent to the information costs) exceeds the other two EVs. The EV for buying is:

$$EV(buy)_t = P(good | e)_t \times G + [1 - P(good | e)_t] \times L + C \times t \quad (12)$$

where G is the value gained for buying a good stock ($G = 80$), L is the value lost for buying a bad stock ($L = -80$), and C is the rating cost ($C = -2$). The EV for rejecting is:

$$EV(reject)_t = C \times t \quad (13)$$

Following Busemeyer and Pleskac (2009) the EV for waiting is:

$$EV(wait)_t = \sum_{j=t}^J \sum_{q=1}^{Q-j+1} P(q)_{j+1} \times O(S_j + c_q)_{j+1} \quad (14)$$

where J is the total number of time points ($J = 6$), Q is the total number of possible ratings at each time point ($Q = 4$), $P(q)$ is the probability of the occurrence of rating q ($P(q) = 0.25$), S is the sum of ratings, and c is the value of rating q ($c = +2, +1, -1, \text{ or } -2$). $O(x)_{j+1}$ refers to the EV of following the optimal policy (to choose the option with the highest EV) at the next step ($j + 1$) given state x . Essentially, Equation 14 requires looking ahead to what states ($x = S_j + c_q$) could occur next, how likely they are going to occur ($P(q)$), and to infer the optimal action at each state and then to recursively evaluate the optimal action at the current state. This procedure has to be repeated for all remaining points of time (from $j = t$ to J) and is therefore most complex at the beginning of each trial.

To fit the model to behavioral data, we used an exponential choice function comparing the EVs of the three choice options:

$$P(choice | t) = \frac{e^{\gamma \times (EV(choice)_t + \delta_{choice})}}{\sum_{choice} e^{\gamma \times (EV(choice)_t + \delta_{choice})}} \quad (15)$$

where γ is a free parameter controlling for the stochasticity of a participant's choices, and δ_{choice} refers to two free parameters for the choice options buy and reject ($\delta_{choice=wait}$ is fixed to 0) allowing overweighting and underweighting of the corresponding choice option (in analogy to asymmetric boundaries for buying and rejecting in SSMs). The exponen-

tial choice rule relates the EVs of the three options (buy, reject, wait) to the probability of buying and rejecting by an S-shaped sigmoid function (i.e., the higher the value for buying compared with the values for rejecting and waiting, the higher the probability of a buy response) and is a standard choice rule for predicting value-based choices (Sutton and Barto, 1998). Note that the optimal solution is a special case of this implementation with $\gamma = \infty$, $\delta_{\text{buy}} = 0$, and $\delta_{\text{reject}} = 0$ (setting γ to ∞ transforms the S-shaped function to a step function so that the option with the highest EV is always chosen, setting δ_{buy} and δ_{reject} to 0 omits any biases in the valuation). Hence, fitting the model to participants' decisions by estimating these free parameters allows accounting for non-optimal behavior in terms of stochastic choices and of overweighting and underweighting of EV(buy) and/or EV(reject) (we were interested in whether participants follow the rationale of the optimal solution in principle, not in whether they solve the task optimally in a strict sense). Note that the probabilities derived from Equation 15 are conditional choice probabilities given t (equivalent to Equations 4 and 5) and have to be implemented in Equation 6 to obtain the actual model predictions.

Model fit and model comparison. We applied maximum likelihood techniques to estimate the model's parameters. The log-likelihood of the data given a model with parameter set Θ for a single participant is:

$$LL_{\text{Model}} = \sum_{n=1}^N \ln[p(\text{buy} | \Theta)_{j,n} \times I_n + P(\text{reject} | \Theta)_{j,n} \times (1 - I_n)] \tag{16}$$

where N is the number of trials ($N = 120$), j refers to the decision point at trial n , and I_n is an indicator function representing whether the participant bought ($I_n = 1$) or rejected ($I_n = 0$) the offer at trial n . The log-likelihood term is used to estimate the deviance $G^2 = -2 \times LL_{\text{Model}}$ (Lewandowsky and Farrell, 2011), which is minimized by finding optimal values for Θ . For optimization procedures, we used the SIMPLEX search method as implemented in the `fminsearch` algorithm in MATLAB (MathWorks). For model comparison the deviance is used for calculating the difference of the Bayesian Information Criterion (BIC) values between a specific model and a reference model as follows:

$$\Delta BIC_{\text{Model}} = 2 \times (LL_{\text{Model}} - LL_{\text{Baseline}}) - k \times \log N \tag{17}$$

where k is the number of free parameters of the specific model. The Baseline model predicts the decision point j at chance level. A priori, chance level is the product of 1 divided by the number of possible decisions (2) and 1 divided by the number of possible decision points (6) (i.e., chance = $1/2 \times 1/6 = 8.3\%$). However, participants did not distribute their responses equally across all choice options and time points; responses at $t = 1$, for instance, were extremely rare ($< 1\%$). To provide a more competitive baseline model we related the chance level to the actual frequency of decisions and decision points $N(\text{choice})_t$, which leads to a stronger model comparison test. The actual chance level for each participant was thus determined by the following:

$$\text{chance} = \sum_{t=1}^J \sum_{\text{choice}} \left[\frac{N(\text{choice})_t}{N} \right]^2 \tag{18}$$

where J refers to the number of possible decision points (i.e., $J = 6$). Our approach yielded an average chance level of 14.2% across all participants (range: 11.4 – 20.2%). Using the individual chance level to calculate LL_{Baseline} , we estimated ΔBIC values for each model in each participant. Note that positive ΔBIC values indicate that the model performs better than the Baseline model (while taking the number of free parameters into account).

Computational models (II). The models and model comparison procedures described above were used to test whether the SSM approach was indeed most suitable to explain participants' choice behavior. As outlined in Results (see Fig. 4A), the time-variant SSM provided by far the closest fit to the data. A separate question, however, refers to the exact nature of the DV, that is, whether people really track the log-evidence or use another quantity. Hence, we set up three additional versions of the time-variant SSM that only differed with respect to the used DV and tested them against the time-variant SSM using the log-evidence (Fig.

4B). The first alternative (*objective* $P(\text{good} | e)$) used the updated probability of a good stock offer $P(\text{good} | e)$, as specified in Equation 1 (more precisely: $DV = P(\text{good} | e)_t - 0.5$, such that the DV is negative, if the probability of a bad stock is higher than for a good stock). For the second alternative (*subjective* $P(\text{good} | e)$), we estimated individual, subjective analogues of $P(\text{good} | e)_t$ based on the probability estimates from the postscan probability estimation task: in each participant, we used a logistic regression to regress the sum of ratings against the probability estimates and used the resulting regression parameters for specifying subjective probabilities in the fMRI task based on the sum of ratings at each time point (Fig. 2, the average regression curve). The third alternative (sum of ratings) simply used the sum of ratings S_t as specified in Equation 3. The models were compared by means of their BIC values.

Statistical analysis of behavioral data. The influence of decision point on reaction times (RTs) and required evidence was examined by one-way repeated-measures ANOVA using only decision points 3, 4, 5, and 6, for which we had data from 25 of 26 participants; decision points 1 and 2 were excluded, as for these we had only data from 6 and 19 participants, respectively. Differences in RTs and required evidence with respect to buy and reject choices were tested by paired-sample t tests separately for each decision point. In addition, we used a 4×2 repeated-measures ANOVA to examine the combined effect of decision point and required evidence on RTs. The factor *required evidence* was created by median-splits of trials with low and high evidence for each decision point in each subject separately. Greenhouse–Geisser correction was used if assumptions of sphericity were violated.

fMRI data acquisition and preprocessing. Whole-brain fMRI data were collected on a 3 T Siemens Trio scanner using a 32-channel head coil. Echo-planar T2*-weighted images (TR 2460 ms, TE 26 ms, FOV 220×220 , flip angle 90°) were acquired using 40 axial slices with a voxel size of $2 \times 2 \times 2$ mm plus a 1 mm gap between slices. Slice orientation was tilted -30° to the anteroposterior commissure axis to reduce signal drop in regions of orbitofrontal cortex (OFC; Deichmann et al., 2003). Additionally, a high-resolution T1-weighted MPRAGE image (voxel size $1 \times 1 \times 1$ mm) was acquired for each subject to improve spatial preprocessing. Preprocessing of fMRI data was performed using SPM8 (Wellcome Trust Center for Neuroimaging, University College London) and commenced with slice timing correction to the middle slice of each volume followed by spatial realignment and unwarping to account for movement artifacts. The individual T1-weighted image was then coregistered to the mean functional image generated during realignment. The coregistered image was segmented into gray matter, white matter, and CSF by the "New Segment" algorithm in SPM8 and the obtained tissue-class images were used to generate individual flow fields and a structural template of all participants by the DARTEL toolbox. Flow fields were used for spatial normalization of functional images to Montreal Neurological Institute space. Images were smoothed by a Gaussian kernel of 6 mm full-width at half-maximum and highpass filtered at 128 s.

Statistical analysis of fMRI data. Statistical analysis comprised four first-level analyses, which were based on the general linear model (GLM) approach as implemented in SPM8. The first GLM examined effects for entire trial durations and therefore included an onset vector for each rating presentation (note that presentation time was at least 2 s to prevent potential nonlinearity in the accumulation of the blood oxygen level-dependent [BOLD] signal that occurs at presentation intervals < 2 s) (Friston et al., 2000). To investigate signals tracking accumulated value, the onset vector was accompanied by the parametric modulator $P(\text{good} | e)_t$, that is, the updated probability of a good offer (parametric modulators are additional regressors in the GLM temporally linked to onset vectors but encoding the modulation of the BOLD signal by a variable of interest like EV or accumulated evidence). To investigate signals tracking accumulated evidence, the onset vector was also accompanied by the parametric modulator $|LE(\text{buy})_t|$, that is, the unsigned log-evidence for buying or rejecting. In addition, we included the time point t (increasing from 1 up to 6 within each trial) as a parametric modulator in this analysis: since evidence tends to increase with time, we ensured that the observed effects cannot be accounted for by a simple linear increase in activity but are specific to the development of evidence. Note that we

omitted the automatic, stepwise orthogonalization of parametric modulators in SPM.

The second and third first-level analyses were set up to further ensure that the accumulation of evidence can be dissociated from a simple linear increase of activity in the brain: the second analysis investigated evidence effects at every time point. There to median-splits for each rating presentation (from 1 to 6) into high and low evidence states were conducted. This resulted in up to 12 separate event-related onset vectors (two evidence states, six time points) depending on the choice profile of each participant (two participants did not have enough trials with a decision at $t = 6$ to create the respective regressors). These onset vectors were entered into a new GLM. Contrast estimates for each regressor were extracted for the peak coordinates of brain regions reported in Figure 7A. The third analysis contrasted specific trials against each other that differed with respect to accumulated evidence. We categorized trials depending on whether the log-evidence (Equation 7) was always positive or always negative throughout the trial (unambiguous) or returning back to zero and/or even switching signs (ambiguous). The rationale of this analysis is that the average unsigned log-evidence $|LE(buy)_i|$ is higher in unambiguous trials, which should induce an earlier onset and higher increase of activation in brain regions that track this log-evidence (Heekeren et al., 2004; Ploran et al., 2007; Ho et al., 2009). Since the decision point tended to be earlier in unambiguous trials, we only took a subsample of all unambiguous and ambiguous trials such that the decision point was exactly matched within each participant (e.g., if there were 10 unambiguous and 20 ambiguous trials with decision point $t = 5$, the 10 unambiguous trials and 10 randomly selected ambiguous trials were selected for the analysis). This ensured that the average trial length did not differ between trial types (unambiguous: 9.79 s, ± 2.00 SD; ambiguous: 9.77, ± 2.00 ; $p = 0.64$). This procedure resulted in 59 trials per participant on average (± 12 SD) that could be analyzed. Two separate onset vectors for the two trial types were created and entered into a new first-level analysis. Time courses were extracted for the peak coordinates of brain regions reported in Figure 7A using the toolbox rfxplot (Gläscher, 2009) for SPM8. Because of the random selection of trials, we repeated this analysis several times with new trial selections to ensure that the selection of specific trials had no substantial influence on the results.

The third GLM examined BOLD responses only at presentation of the first rating. The corresponding onset vector was accompanied by parametric modulators for $P(\text{good} | e)_p$, $|LE(buy)_i|$, and for the cumulative sum of the probability (CSP) to respond until the real decision point, j , as given by:

$$CSP = \sum_{i=1}^j P(buy)_i + P(reject)_i \quad (19)$$

according to the best performing model (the time-variant SSM). CSP represents the probability that a response has been made at the decision point or earlier. If this value is low in a specific trial, then the response was made relatively early (compared with what the behavioral model predicted based on the average choice profile of a participant). If this value is high, then the response was made relatively late. In other words, CSP quantifies the difference between behavior in a specific trial and the average behavior of a participant (as captured by the behavioral model). We used this regressor to track for regions in the brain that predict (at trial start) trial-by-trial fluctuations in the tendency to respond earlier than on average (Fig. 1C). The other two parametric modulators were used to exclude confounding effects when looking at the third modulator (CSP). Importantly, we excluded trials with decision point 1 for this analysis. All GLM analyses also included an onset vector for the response together with a parametric modulator coding for the specific response (buy vs reject) that was also used to create response-related regions of interest (ROI, see below).

At the group-level, we used the full factorial design as implemented in SPM8 (controlling for nonsphericity of the error term) to test for effects related to accumulated value, accumulated evidence, and trial-by-trial deviations (CSP). To test for the functional criterion of creating response-related ROIs, first-level contrast images of the specific response (buy vs reject) were entered into a regression analysis: the regression

contrasted participants, who bought with their right and rejected with their left index finger with those participants who bought with their left and rejected with their right index finger. The statistical threshold for imaging results was set to $p < 0.05$, familywise error rate (FWE) corrected for small volumes. Small volumes were either spherical search volumes (sphere radius: 10 mm) around peak coordinates from previous studies that tested for comparable effects of interests or anatomical masks derived from MRI atlases (Lancaster et al., 2000; Tzourio-Mazoyer et al., 2002). Center coordinates for spherical search volumes were $[x = -3, y = 42, z = -6]$ for vmPFC (Chib et al., 2009), $[\pm 42, 33, 9]$ for OFC (Gläscher et al., 2009), and $[\pm 14, 10, -10]$ for VS (O'Doherty et al., 2004). Anatomical masks were created for pre-SMA, insula, and caudate nucleus. ROIs for response-related areas were generated based on two criteria: the functional criterion was that activation related to the modulator of the specific response (buy vs reject) exceeded a threshold of $p < 0.05$ FWE-corrected for whole-brain, and the anatomical criterion was that activation occurred within an atlas mask of the precentral gyrus (Tzourio-Mazoyer et al., 2002). Contrast estimates for value-related activation in these ROIs were derived using the toolbox rfxplot (Gläscher, 2009) for SPM8. Other regions were reported if they survived a threshold of $p < 0.05$, FWE-corrected for whole-brain. For display purposes, we used a threshold of $p < 0.001$ (uncorrected) with 10 contiguous voxels. Activations are depicted on a skull-stripped overlay of the mean structural T1-weighted image from all 26 participants.

Results

Behavioral results

The final sample included $n = 26$ participants. All participants collected a positive amount of on average 960 points (± 528 SD, range = 172–2288) indicating that they understood the task and used the ratings to guide their choices (random choices would inevitably lead to a negative score). However, a comparable sample of 26 optimal agents would have collected a significantly higher amount of on average 1239 points (± 417 SD, 556–2074; $t_{(25)} = 3.65$; $p < 0.001$). The average decision point (i.e., the rating at which a response was given) was 4.16 (± 0.70 SD, 2.80–5.32) as compared with the optimal solution with an average decision point of 3.67. The majority of 19 of the 26 participants responded on average later than the optimal decision point (sign test: $p = 0.029$). The mean decision point for each participant was negatively correlated with self-evaluated novelty seeking ($r = -0.46$, $p = 0.017$), that is, more impulsive participants required less evidence to make their decisions (Fig. 3A). There was no significant correlation with the other queried scales of the TCI questionnaire (harm avoidance: $r = 0.23$, $p = 0.252$; cooperative-ness: $r = -0.26$, $p = 0.205$).

In general, evidence required for deciding decreased linearly with increasing points of time ($F_{(3,72)} = 42.67$, $p < 0.001$) similarly for buy and reject decisions (Fig. 3B). This result is in line with a time-variant SSM that assumes decreasing rather than fixed decision boundaries: even if the accumulated evidence remains at a low level, the decreasing boundaries ensure that the decision process does not last forever (Ditterich, 2006a,b; Churchland et al., 2008; Cisek et al., 2009). Differences between buy and reject decisions were restricted to the last decision point, for which reject decisions required less evidence ($t_{(23)} = 3.25$, $p = 0.003$). This effect appears to be driven by the fact that participants rather rejected than bought the stocks, if there was no evidence for either choice option at this point ($t_{(23)} = 2.38$, $p = 0.026$). RTs also decreased with time ($F_{(3,72)} = 4.70$, $p = 0.011$), again similarly for buy and reject decisions (Fig. 3C). Buy decisions were faster than reject decisions for decision point 3 ($t_{(23)} = 3.17$, $p = 0.004$) but not for the remaining time points. To examine the interplay of required evidence, decision points, and RTs, we conducted a median-split of trials for each decision point to separate decisions

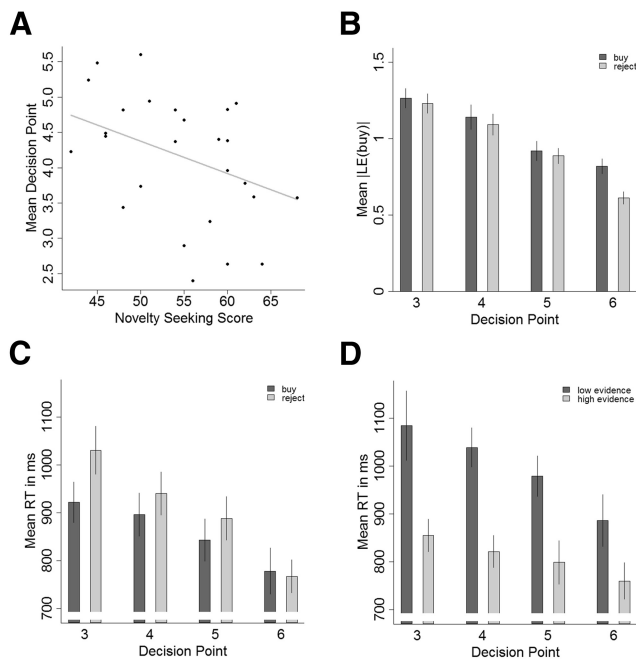


Figure 3. *A*, Negative relationship between self-evaluated novelty seeking and the mean decision point per participant. *B*, Evidence at the decision point decreased linearly with time similarly for buy and reject decisions, suggesting a time-variant accumulation process. *C*, As for evidence, RTs were reduced at later decision points similarly for buy and reject decisions. *D*, Combined effects of evidence and time point on RTs: higher evidence and elapsed time independently reduced reaction times. Only data for decision points 3–6 are shown as there were too few decisions made at $t = 1$ and 2.

at high and low evidence. Figure 3*D* shows that within each decision point, decisions were made faster when evidence was high ($F_{(1,24)} = 69.30, p < 0.001$) and, as already mentioned, RTs decreased with later decision points ($F_{(3,72)} = 4.52, p = 0.013$). Although we did not model RTs within each decision point (only decision points per se), the observed patterns nicely fit into the logic of a time-variant accumulation process: RTs depend on the distance of the decision variable from the boundaries and on the rate at which this distance is surmounted. Thus, both higher evidence and lower boundaries reduce RTs.

Model comparison

Given the behavioral results, it is not surprising that our model comparison also revealed the time-variant SSM to be most adequate for predicting decisions and decision points. Figure 4*A* shows the comparison of different computational models with a baseline model by means of their average BIC values (see Materials and Methods). The current evidence model, which relies only on the currently presented rating (without evidence accumulation) performed worst, even worse than the Baseline model. The fix variance model and the optimal solution both provided a better fit but still were inferior to the three SSM approaches. Within the SSM models the SSM with forgetting performed worse than the standard SSM, but the time-variant SSM was clearly superior to both of them. For 25 of 26 participants the time-variant SSM provided the best fit (for the remaining participant, the SSM with forgetting was best). In addition, the time-variant SSM appears to describe the data well in absolute terms. If we take, for instance, the highest probability of a specific response at every time point according to this model and compare it to the actual decisions and decision points, we see that the model made correct predictions in 62.6% (± 9.7 SD) of the trials, which is well

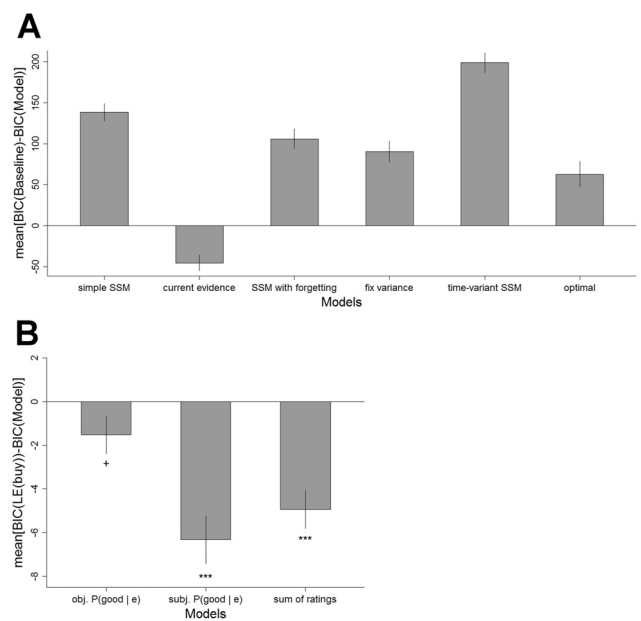


Figure 4. Model comparisons. *A*, Comparison of SSM approaches and other computational models suggested the time-variant SSM to provide the best description of choice behavior (depicted are the mean differences between the BIC value of each model and the Baseline model; positive values indicate better model fit). *B*, Comparison of time-variant SSMs with different DVs (depicted are the mean differences between the BIC value of each alternative and the model that used the log-evidence as DV). $^+p < 0.10$; $***p < 0.001$.

above the estimated chance level of 14.2% (± 2.3 SD). Figure 5 shows that frequencies of buy and reject decisions at every time point are accurately recovered by the time-variant SSM for all participants. Furthermore, estimates for the free parameters of this model accurately reflected interindividual differences in behavior: the mean decision point correlated with the height of the two boundaries (for θ_{buy} : $r = 0.57, p = 0.002$; for $|\theta_{\text{reject}}|$: $r = 0.58, p = 0.001$), showing that higher boundaries imply later decisions. Also, the difference between the height of the two boundaries $\theta_{\text{buy}} - |\theta_{\text{reject}}|$ correlated positively with the number of reject decisions ($r = 0.83, p < 0.001$), showing that asymmetric boundaries promote the preference of a particular choice. If θ_{reject} is small (large) compared with θ_{buy} , the accumulation process is more likely to cross θ_{reject} (θ_{buy}), and therefore more offers are rejected (bought). The parameter λ , which should model the strength in the decrease of the decision boundaries, was correlated with interindividual differences in the strength of the effect reported in Figure 3*B* ($r = 0.71, p < 0.001$).

Having shown that a time-variant SSM is most adequate for predicting decisions and decision points in our task, we tested whether a DV other than the log-evidence for buying ($LE(\text{buy})_t$) might provide an even closer fit to the data (within the framework of the time-variant SSM). Note that an exact estimation of $LE(\text{buy})_t$ is very complex (see Materials and Methods) and our participants might instead use a heuristic approximation to it, such as the sum of ratings (see Equation 3), as DV. We thus compared different versions of the time-variant SSM that differed only with respect to the underlying DV. Figure 4*B* shows that the difference in the average BIC values between the original version and the three alternatives was always in favor of the log-evidence model, though there was only a non-significant trend ($p = 0.089$) for the comparison with the objective $P(\text{good} | e)$ model. Since all DVs were highly correlated to

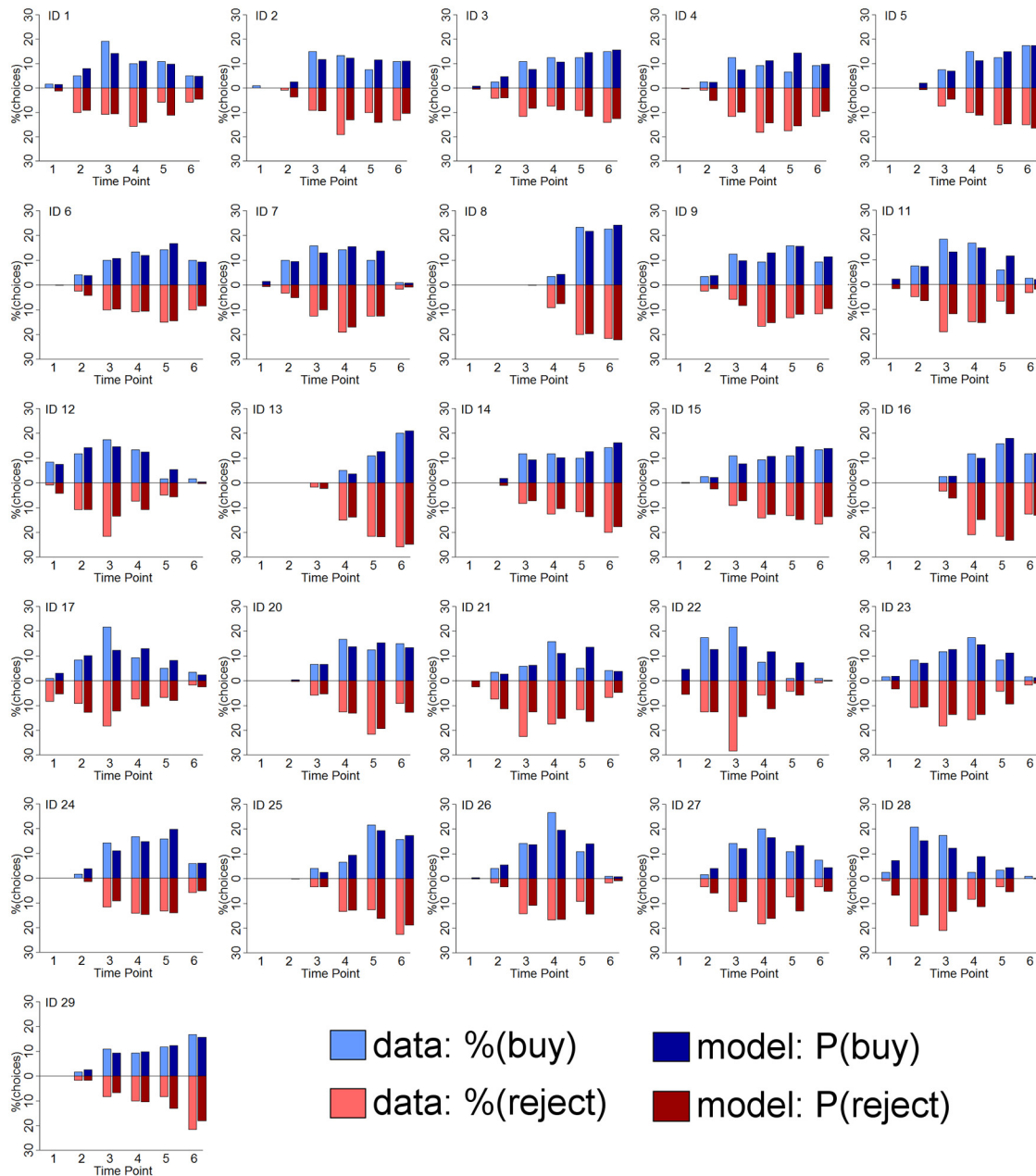


Figure 5. Individual decision profiles and model predictions. For each participant, the average percentages of buy and reject decision at each time point is shown along with the average probabilities for buy and reject decisions according to the time-variant SSM. Note that the model is able to capture the variability in decision points and the differences in the tendency to rather buy or reject across participants.

each other (all $r < 0.9$), however, we did not expect to find large differences.

Brain regions tracking accumulated value

In our paradigm, the EV essentially varies with changes in the probability of a good stock ($P(good | e)_t$), when new ratings are disclosed, and we assumed participants to track this updating process. Results from the probability estimation task (after scanning) suggested that our participants were able to approximate $P(good | e)_t$ with a tendency to overestimate low and underestimate high probabilities (Fig. 2). For the fMRI analysis, $P(good | e)_t$ was implemented as a parametric modulator at each rating presentation. In line with our hypothesis, we found this variable to be significantly correlated with BOLD signals in the vmPFC and the right VS. Further regions showing a positive relationship with

accumulated value were the left and right OFC (extending into the anterior insula) and the caudate nucleus (Fig. 6; see also Table 1 for all fMRI results).

Brain regions tracking accumulated evidence

Next, we identified brain regions associated with accumulation of evidence in general, that is, regardless of for buying or for rejecting. The unsigned accumulated log-evidence for buying or for rejecting the current offer ($|LE(buy)_t|$) was used as the respective parametric modulator. Since evidence tends to increase with time, we also included a regressor modeling a linear increase within each trial. This allowed us to separate effects of evidence from any other phenomena that might constantly increase with time (e.g., representation of information costs, working memory demands, response urgency). Accumulated evidence was associ-

ated with BOLD signals in the insula, pre-SMA, caudate nucleus, and right dorsolateral prefrontal cortex (dlPFC; Fig. 7A).

We verified that activation in the located areas indeed depended on evidence (beyond a simple linear increase) by two additional analyses. First, we conducted median-splits for every time point within trials (from 1 to 6) separating high from low states of evidence and extracted the average BOLD signal in the located areas. As evident from Figure 7B, activation in these regions increased with time but was also higher when evidence was high. Second, we compared the average BOLD time series in these regions of unambiguous trials, in which evidence was always positive or always negative, against ambiguous trials, in which evidence returned to zero or even switched signs. Note that trials were selected such that decision points were matched between conditions to avoid differences in trial length (see Materials and Methods). Because the average ($|LE(buy)_i|$) is higher in unambiguous compared with ambiguous trials (unambiguous: 0.82, ± 0.10 ; ambiguous: 0.44, ± 0.06 ; $t_{(23)} = 31.13$, $p < 0.001$), we expected to find an earlier and stronger increase of activity (Heekeren et al., 2004; Ploran et al., 2007; Ho et al., 2009). Figure 7C shows that this pattern was matched for the regions reported in Figure 7A (we refrain from reporting statistics for these additional analyses as they are not independent from the parametric analysis; Kriegeskorte et al., 2009). The regressor of a linear increase revealed a large amount of brain areas, including the regions correlating with accumulated evidence (data not shown). Hence we were able to isolate regions, which specifically tracked the accumulation of evidence, from a broader network of brain areas with a linearly increasing activation pattern.

Brain regions accounting for trial-by-trial variability in behavior

Neuroimaging studies indicate that increased baseline activity in pre-SMA and caudate nucleus at the start of the accumulation process mediate faster responses in perceptual decisions under time pressure (Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008). This effect presumably refers to a reduced distance to the decision boundary (Bogacz et al., 2010). Accordingly, we attempted to identify fMRI BOLD responses at the beginning of each trial (i.e., at presentation of the first rating), which predicted trial-by-trial variability in response tendencies (i.e., whether the response was given earlier or later than on average as suggested by the behavioral model). To do so, we included the CSP (Equation 19) according to the best performing model at the decision point as an event-related parametric modulator at the onset of each trial. The rationale of this analysis is illustrated in Figure 1C: the behavioral model predicts the decision point by a probability mass function. The cumulative sum of this function until t refers to the probability that a decision has been made until t and CSP is this cumulative sum at the real decision point. If CSP is high, then the decision has been made relatively late (because the model suggested an earlier response with high probability). If CSP is low, then the decision has been made comparatively early (because the model suggested a later response with high probability). We tested for brain regions being negatively associated with CSP at trial start, such that activity in these regions predicted comparatively early responses. Note that by identifying brain regions associated with a parametric modulator coding this variable we can account for trial-by-trial

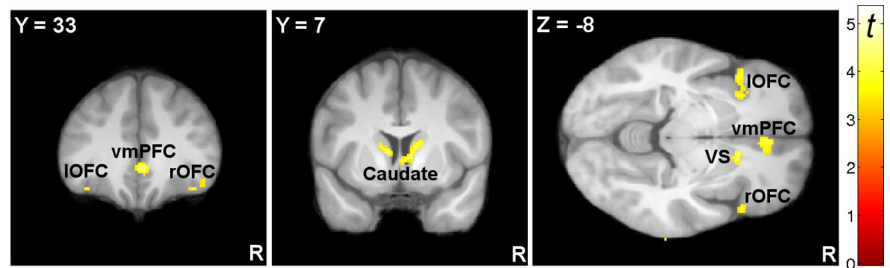


Figure 6. The fMRI BOLD signals in vmPFC, left and right OFC, VS, and caudate nucleus were correlated with the regressor for accumulated value.

Table 1. Peak coordinates and statistics of fMRI analyses

Contrast	Name of region	MNI coordinates (mm)			Statistics		
		x	y	z	t value	Z value	p value
Accumulated value	vmPFC	2	36	-4	4.88	4.39	0.002
	Left OFC	-38	32	-18	4.06	3.75	0.015
	Right OFC	40	36	-18	4.08	3.77	0.014
	VS	12	18	-8	4.33	3.97	0.007
	Left caudate nucleus	-8	8	8	4.25	3.91	0.014
	Right caudate nucleus	8	6	4	4.47	4.08	0.008
Accumulated evidence	Left insula	-28	24	-2	8.38	7.02	<0.001
	Right insula	36	22	4	8.99	7.38	<0.001
	Left pre-SMA	-6	8	46	5.07	4.68	0.001
	Right pre-SMA	8	8	54	6.09	5.47	<0.001
	Left caudate nucleus	-12	8	6	5.31	4.88	<0.001
	Right caudate nucleus	10	8	10	4.60	4.31	0.003
	dlPFC	44	6	28	6.39	5.69	0.001
	Left thalamus	-4	-18	-2	5.90	5.33	0.006
	Right thalamus	6	-24	-2	6.37	5.67	0.001
	Precuneus	20	-60	28	5.75	5.22	0.010
Postcentral gyrus	-46	-26	46	5.74	5.21	0.010	
Inverse CSP	Insula	-28	24	8	4.26	4.02	0.016
	Left pre-SMA	-10	10	46	4.15	3.92	0.023
	Right pre-SMA	6	10	56	4.00	3.80	0.038
	Caudate nucleus	8	8	-2	3.87	3.68	0.028
	IPS	32	-52	48	6.30	5.63	0.001
	Midbrain	-8	-22	-12	6.48	5.75	0.001

dlPFC, dorsolateral prefrontal cortex; IPS, intraparietal sulcus; pre-SMA, presupplementary motor area; MNI, Montreal Neurological Institute; OFC, orbitofrontal cortex; vmPFC, ventromedial prefrontal cortex; VS, ventral striatum. Small volume corrected regions (vmPFC, OFC, VS, caudate nucleus, insula, pre-SMA) are listed first followed by other regions that survived a threshold of $p < 0.05$ FWE-corrected at whole brain.

variability in response tendencies that are inherently inexplicable by the behavioral models, as we derived this regressor from their “inability” to perfectly predict the decision point.

When we tested for brain regions predicting early responses in this way, we obtained a similar pattern of activity as for the accumulation of evidence: significant activations were found in pre-SMA, right caudate nucleus, and left insula (right insula activation was also observed but did not reach significance), as well as in the right intraparietal sulcus (IPS) (Fig. 7D). Note that we controlled for effects of evidence, value, and the response hand by including the respective regressors in the analysis. Furthermore, trials with decisions made at $t = 1$ were excluded from this analysis to avoid confusion with response-related signals.

Value correlates in motor areas

A continuous transmission of decision-related information to response execution units in the brain has been proposed by many proponents of SSMs (Gold and Shadlen, 2007; Heekeren et al., 2008; Cisek and Kalaska, 2010). We therefore tested whether value information is continuously and immediately

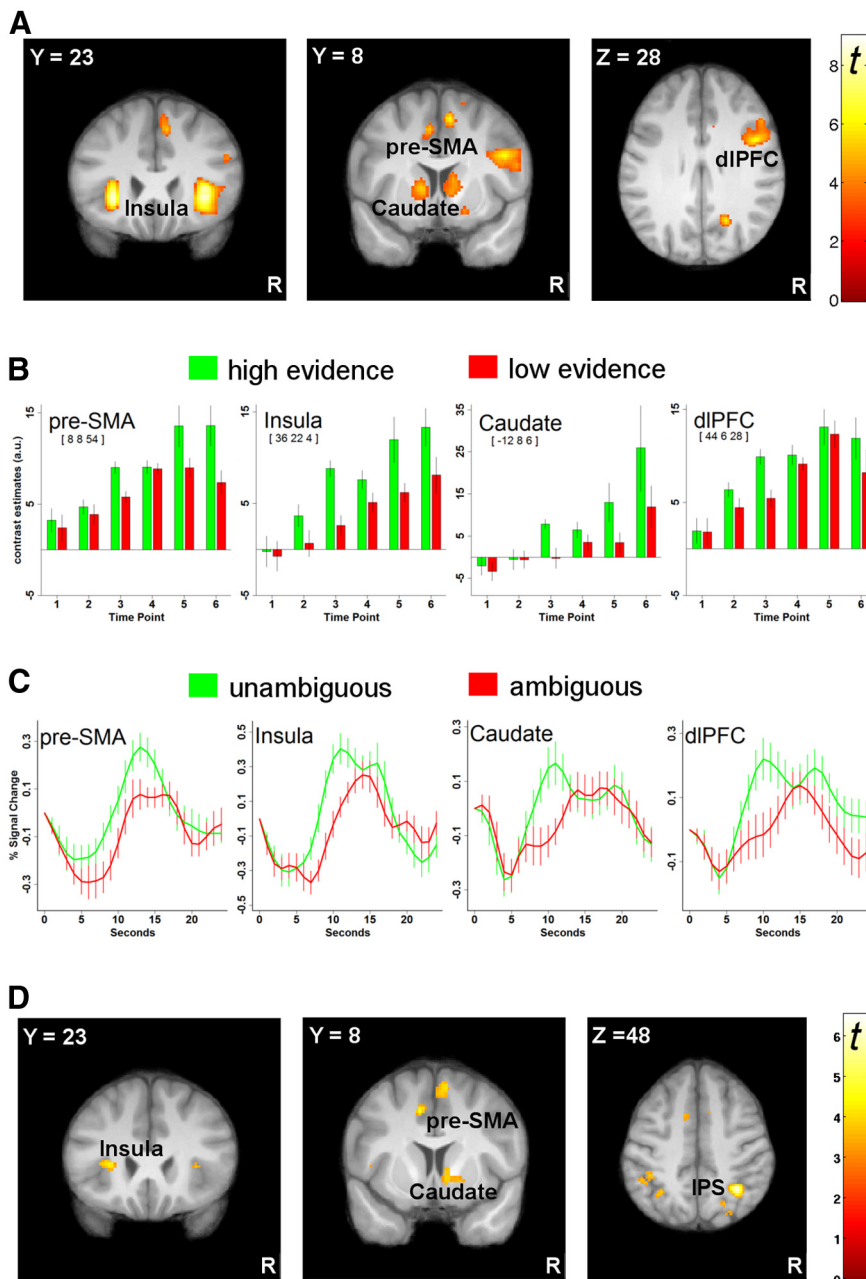


Figure 7. *A*, The fMRI BOLD signals in anterior insula, pre-SMA, caudate nucleus, and dlPFC were correlated with the regressor for accumulated evidence. *B*, Contrast estimates for the regions reported in *A* at every time point separated for high (green) and low (red) states of evidence. Signals increased with time but were also modulated by the amount of evidence at each time point. *C*, fMRI time series for the regions reported in *A* separated for unambiguous (green) and ambiguous (red) trials. Unambiguous trials lead to an earlier and stronger increase of activity. *D*, Activity at trial starts in many regions related to accumulated evidence also predicted trial-by-trial deviations of actual behavior from model estimates: increased activity in the insula, pre-SMA, and caudate nucleus as well as IPS was negatively associated with CSP, indicating that higher signals in these areas preceded decisions that were made earlier than on average (see Fig. 1C to understand how CSP was created).

transferred to response-related brain regions, as has been shown in the monkey (Pastor-Bernier and Cisek, 2011) and in humans for perceptual decisions (Donner et al., 2009). First, we identified response-related cortical areas as ROIs including only those voxels that correlated with the event-related regressor of the specific response (buy or reject) at $p < 0.05$ (FWE-corrected at whole brain). In addition, we restricted the ROIs anatomically by excluding regions not belonging to the precentral gyrus (such as the basal ganglia, which are known to encode both

response- and value-related information). Figure 8*A* depicts the derived ROIs. To test for effects of value, we then extracted the parameter estimates for $P(\text{good} | e)_t$ separately for each ROI in each hemisphere. The rationale is that activity in the ROI contralateral to the response hand for buying should be positively associated with $P(\text{good} | e)_t$, whereas activity in the ROI contralateral to the response hand for rejecting should be negatively associated. This approach is reminiscent of the lateralized readiness potential in electrophysiology (Coles, 1989). As the assignment of response hand (left, right) to choices (buy, reject) was counterbalanced across participants, we could test for the effect of value on lateralized activation of ROIs using a 2×2 ANOVA with response to hand assignment as a between-subject factor and hemisphere of ROI as a within-subject factor. As can be seen from Figure 8*B* (left), there was a strong interaction effect indicating a positive relationship with $P(\text{good} | e)_t$ in the ROI for buying and a negative relationship in the ROI for rejecting ($F_{(1,24)} = 51.84; p < 0.001$).

The interaction effect supports the hypothesis that value information is continuously transmitted to the motor system. As the parameter $P(\text{good} | e)_t$ is defined at every rating presentation, however, one might argue that the observed lateralization could solely originate from response preparation directly preceding the response. In other words, at the point when a stock is finally bought, $P(\text{good} | e)_t$ is usually high and thus correlated with the response itself. We therefore performed a more rigorous analysis, which was restricted to the presentation of the first rating. This time point did not coincide with responses themselves: the few trials with a response at this time point ($<1\%$) were excluded from the analysis. As evident from Figure 8*B* (middle), we could replicate the interaction of response hand assignment and hemisphere for presentation of the first rating only ($F_{(1,24)} = 22.74; p < 0.001$). To further strengthen our argument, we separated first ratings into trials of high and low evidence: A “+ +” (“- -”) at $t = 1$ provides higher evidence for buying (rejecting) than a “+”

(“-”). We expected the lateralization effect to be modulated by the amount of evidence. This was indeed the case, as revealed by a significant three-way interaction of response hand, hemisphere, and evidence ($F_{(1,24)} = 9.12; p = 0.006$; the interaction of response hand and hemisphere remained significant: $F_{(1,24)} = 15.49; p < 0.001$). Figure 8*B* (right) illustrates that the lateralization effect was more pronounced when evidence for or against buying was high as compared with low.

Interestingly, this early lateralization effect is related to the average decision point per participant: we obtained a lateralization score for each participant by subtracting the parameter estimates of the ROIs contralateral to the response hands for buying and rejecting from each other, such that higher scores refer to stronger lateralization. This score showed a strong negative correlation with the mean decision point ($r = -.62; p < 0.001$) (Fig. 8C), indicating that the influence of the first rating's value on motor execution areas was greater for participants who responded earlier on average.

Discussion

In this study, we investigated the process underlying the temporal evolution of value-based decisions in the human brain, using fMRI in combination with a cognitive modeling approach. The behavioral data were best described by a time-variant sequential sampling model that accounts for the linear decrease of required evidence with time and is compatible with the similar decrease in RTs. Interestingly, we found that more impulsive participants responded earlier, lending support for the external validity of our paradigm. The fMRI results suggest that the brain integrates multiple sources of information by forming an updated value representation in dopaminergic areas including vmPFC and VS. BOLD signals in the anterior insula, pre-SMA, and caudate nucleus were related to the accumulation of evidence regardless of the specific response or response hand. Activity in these areas at the onset of a trial also predicted later deviations from behavioral models. Finally, we detected value-related information in motor execution areas within the precentral gyrus already at the onset of the accumulation process; the magnitude of this effect was modulated by interindividual differences in the height of the decision boundaries.

In the time-variant SSM, evidence is accumulated until one of the decision boundaries has been reached, but these boundaries decrease over time implying that less and less evidence is needed for committing a decision. This result is in line with recent evidence from recording and modeling neurophysiological data in monkeys (Ditterich, 2006a,b; Churchland et al., 2008) and behavioral data in humans (Cisek et al., 2009). Notably, our paradigm bears particular similarities to the task of Cisek et al. (2009) (limited amount of information; objectively changing evidence within trials). These characteristics most likely led to the observed behavioral effects and as a consequence to the superior fit of the time-variant model. If the specific nature of the environment shapes the decision-making process, it becomes important to characterize the essential features of the decision problem (Simon, 1956; Payne et al., 1988; Todd and Gigerenzer, 2007). We note that value-based decisions are usually made in situations with a restricted amount of information and time (e.g., booking hotels on the Internet, shopping in department stores) as

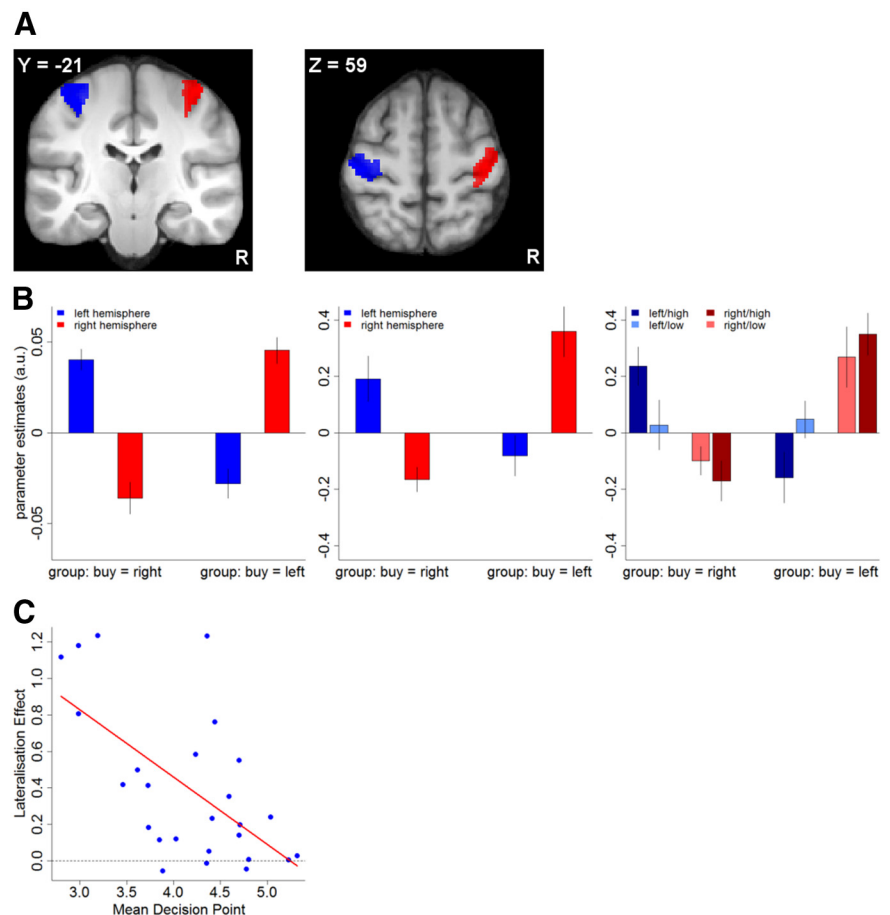


Figure 8. Value signals in motor execution areas. **A**, ROIs in precentral gyrus based on response-specific activity. **B**, Parameter estimates for BOLD signals in ROIs related to accumulated value at all rating presentations (left) and at presentation of the first rating only (middle and right), separated by groups (participants who bought with right/left hand and rejected with left/right hand). The right panel further illustrates the modulatory effect of evidence by separating first ratings with high evidence for/against buying (dark blue/red) from low evidence ratings (light blue/red). **C**, The lateralization effect at the first rating (**B**, middle) correlated negatively with the mean decision point.

well as with highly conflicting information (e.g., divergent hotel reviews, low prices but bad quality) just as it was implemented in the current experimental design. Computational models that account for the need to terminate decisions at some point seem therefore particularly attractive in the framework of economic choices.

Updating the probability of a good stock or the log-evidence for buying (which we used for modeling behavior) requires complex calculations and it appears unrealistic that humans are capable of exactly estimating them. Therefore we tested alternative ways of forming subjective beliefs, such as a heuristic that only counts the number of ratings (Anderson and Holt, 1997) or a nonlinear function of this sum. However, our modeling results do not support these alternatives. Interestingly, Soltani and Wang (2010) suggested that probabilistic inferences between arbitrary stimuli and rewards can be learned through feedback-related changes in synaptic connectivity, which allows good approximations to the mathematical quantities (Engel and Singer, 2008). Although such learning processes cannot fully account for our participants' behavior (feedback was only provided in a few training trials, in which performance was already quite high), future studies should investigate how humans combine instructed information and feedback learning to form a DV when making sequential decisions. The present study did not attempt to make a strong claim about the exact nature of the DV.

For perceptual decisions, pre-SMA and the caudate nucleus have been linked to an adaptive adjustment of the decision threshold by an increased baseline activity under heightened time pressure (Lo and Wang, 2006; Forstmann et al., 2008; Ivanoff et al., 2008; van Veen et al., 2008; Bogacz et al., 2010). Our finding that comparatively early responses are predicted by increased BOLD signals in these areas at trial start is in accordance with the baseline hypothesis and extends it to the domain of value-based decisions. We did not introduce different conditions of time pressure but captured trial-by-trial fluctuations in response tendencies via fMRI that remained unexplained by the behavioral model. We suggest that these fluctuations reflected the probabilistic nature of value-based decisions (Mosteller and Nogee, 1951; Rieskamp, 2008). Future research should test whether voluntary adjustments and involuntary fluctuations of the decision threshold in value-based decisions recruit overlapping brain circuits as suggested for perceptual decisions (van Maanen et al., 2011).

Additionally, activity in the same areas (pre-SMA and caudate nucleus) steadily increased during trials and also tracked accumulated evidence (regardless of for or against accepting the offer). The constant signal increase might be related to the decrease in the decision threshold with time (increased rather than decreased neuronal signals have typically been suggested to reflect reduced decision thresholds) (Ditterich, 2006a,b; Churchland et al., 2008; Cisek et al., 2009; Bogacz et al., 2010), though other explanations seem possible. Altogether, the activation patterns observed in pre-SMA and caudate nucleus tightly mimic all model-derived variables that contribute to predicting the decision point (trial-by-trial fluctuations, decrease in decision thresholds, accumulated evidence). We therefore suggest that these regions signal the general willingness to respond as it evolves throughout the decision-making process. This view is in line with research on volition indicating that a gradual buildup of activity in the pre-SMA precedes self-generated actions (Libet et al., 1983; Haggard, 2008; Fried et al., 2011) and predicts when but not which decisions are made (Soon et al., 2008). The insula, showing the same effects, has been linked to evidence accumulation in perceptual decision making independent of response modalities (Ho et al., 2009). Fluctuations in insula activity might also be related to variability in risk-seeking tendencies (to respond earlier also means to behave more risk seeking) (Preusschoff et al., 2008; Mohr et al., 2010).

It has been repeatedly shown that the vmPFC and the VS parametrically encode the EV of various objects and offers (Yacubian et al., 2006; Chib et al., 2009; FitzGerald et al., 2009; Lebreton et al., 2009). In line with this and with the notion of a working memory system for rewards (Wallis, 2007), we show that these regions maintain and update the representation of value over time during sequential decision making, as activity followed the Bayesian updating rule for $P(\text{good} | e)_t$ (although we do not claim that the brain is capable of integrating value information in an optimal way). Interestingly, the caudate nucleus was the only region linked to EV (Fig. 6) as well as to evidence accumulation in general (Fig. 7). This finding is in accordance with a recent study in nonhuman primates, which proposed the caudate nucleus to encode multiple computations during perceptual decisions (Ding and Gold, 2010; Cai et al., 2011). Thus, the caudate nucleus appears to play a major role in decision making as it integrates various signals relevant for the timing and the specificity of decisions.

The ongoing representation of values was also observed in motor execution areas: activity in the primary motor cortex contralateral to the response hand used for buying was positively

associated with the $P(\text{good} | e)_t$ regressor, and the opposite effect was observed for the side contralateral to the response hand for rejecting. Further analyses revealed that this lateralization occurred already at trial start (although responses were made >10 s later on average) and was modulated by the amount of evidence, suggesting an immediate, continuous, and fine-grained representation of value in the brain's motor system. In contrast to the effects found in vmPFC and VS, this effector-specific encoding of value was influenced by how early participants responded on average. Hence, we suggest that the motor output system combined information signaling the general tendency to respond, as encoded in regions like the pre-SMA, with information signaling the specificity of the response (buy or reject), as provided by regions like the vmPFC.

Previous studies using sequential decision-making paradigms (Yang and Shadlen, 2007; de Lange et al., 2010; Stern et al., 2010) attempted to dissociate signals related to the accumulation of evidence from preparatory motor activity by prohibiting participants to respond before all information had been presented. In contrast, in our study participants were allowed to make their decision at any point. By allowing the decision time to vary we were able to use SSMS to explain the varying decision points and to examine trial-by-trial variability as outlined above. Furthermore, restricting the decision point to the end of the accumulation phase does not necessarily prevent preparatory motor activity to occur during accumulation (Donner et al., 2009).

In conclusion, our data show that dopaminergic areas translate perceptual input into a sustained representation of value that is flexible with respect to newly incoming information. Motor preparation circuits signal the willingness to respond, which is affected by the distinctiveness of value information but also by stochastic fluctuations. Output areas do not simply retrieve the completed decision but reflect the development of the accumulation process right from the beginning.

References

- Anderson LR, Holt CA (1997) Information cascades in the laboratory. *Am Econ Rev* 87:847–862.
- Bogacz R, Wagenmakers EJ, Forstmann BU, Nieuwenhuis S (2010) The neural basis of the speed-accuracy tradeoff. *Trends Neurosci* 33:10–16.
- Busemeyer JR, Pleskac T (2009) Theoretical tools for understanding and aiding dynamic decision making. *J Math Psychol* 53:126–138.
- Busemeyer JR, Townsend JT (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol Rev* 100:432–459.
- Cai X, Kim S, Lee D (2011) Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 69:170–182.
- Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci* 29:12315–12320.
- Churchland AK, Kiani R, Shadlen MN (2008) Decision-making with multiple alternatives. *Nat Neurosci* 11:693–702.
- Cisek P, Kalaska JF (2010) Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci* 33:269–298.
- Cisek P, Puskas GA, El-Murr S (2009) Decisions in changing conditions: the urgency-gating model. *J Neurosci* 29:11560–11571.
- Cloninger CR (1994) The temperament and character inventory (TCI): a guide to its development and use. St. Louis: Center for Psychobiology of Personality, Washington University.
- Coles MG (1989) Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology* 26:251–269.
- Cox DR, Miller HD (1965) The theory of stochastic processes. London: Chapman and Hall.
- Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19:430–441.
- de Lange FP, Jensen O, Dehaene S (2010) Accumulation of evidence during

- sequential decision making: the importance of top-down factors. *J Neurosci* 30:731–738.
- Ding L, Gold JI (2010) Caudate encodes multiple computations for perceptual decisions. *J Neurosci* 30:15747–15759.
- Ditterich J (2006a) Stochastic models of decisions about motion direction: behavior and physiology. *Neural Netw* 19:981–1012.
- Ditterich J (2006b) Evidence for time-variant decision making. *Eur J Neurosci* 24:3628–3641.
- Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol* 19:1581–1585.
- Engel C, Singer W (2008) Better than conscious?: decision making, the human mind, and implications for institutions. Cambridge, MA: MIT.
- FitzGerald TH, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29:8388–8395.
- Forstmann BU, Dutilh G, Brown S, Neumann J, von Cramon DY, Ridderinkhof KR, Wagenmakers EJ (2008) Striatum and pre-SMA facilitate decision-making under time pressure. *Proc Natl Acad Sci U S A* 105:17538–17542.
- Fried I, Mukamel R, Kreiman G (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69:548–562.
- Friston KJ, Mechelli A, Turner R, Price CJ (2000) Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *Neuroimage* 12:466–477.
- Gläscher J (2009) Visualization of group inference data in functional neuroimaging. *Neuroinformatics* 7:73–82.
- Gläscher J, Hampton AN, O'Doherty JP (2009) Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cereb Cortex* 19:483–495.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
- Haggard P (2008) Human volition: towards a neuroscience of will. *Nat Rev Neurosci* 9:934–946.
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.
- Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467–479.
- Ho TC, Brown S, Serences JT (2009) Domain general mechanisms of perceptual decision making in human cortex. *J Neurosci* 29:8675–8687.
- Ivanoff J, Branning P, Marois R (2008) fMRI evidence for a dual process account of the speed-accuracy tradeoff in decision-making. *PLoS ONE* 3:e2635.
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540.
- Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT (2000) Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 10:120–131.
- Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64:431–439.
- Lewandowsky S, Farrell S (2011) Computational modeling in cognition: principles and practice. Thousand Oaks, CA: Sage.
- Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106:623–642.
- Lo C-C, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci* 9:956–963.
- Mohr PN, Biele G, Heekeren HR (2010) Neural processing of risk. *J Neurosci* 30:6613–6619.
- Mosteller F, Noguee P (1951) An experimental measurement of utility. *J Polit Econ* 59:371–404.
- 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.
- Pastor-Bernier A, Cisek P (2011) Neural correlates of biased competition in premotor cortex. *J Neurosci* 31:7083–7088.
- Payne JW, Bettman JR, Johnson EJ (1988) Adaptive strategy selection in decision making. *J Exp Psychol Learn Mem Cogn* 14:534–552.
- 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.
- Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
- Ploran EJ, Nelson SM, Velanova K, Donaldson DI, Petersen SE, Wheeler ME (2007) Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J Neurosci* 27:11912–11924.
- Preusschoff K, Quartz SR, Bossaerts P (2008) Human insula activation reflects risk prediction errors as well as risk. *J Neurosci* 28:2745–2752.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59–108.
- Rieskamp J (2008) The probabilistic nature of preferential choice. *J Exp Psychol Learn Mem Cogn* 34:1446–1465.
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475–9489.
- Simon HA (1956) Rational choice and the structure of the environment. *Psychol Rev* 63:129–138.
- Soltani A, Wang XJ (2010) Synaptic computation underlying probabilistic inference. *Nat Neurosci* 13:112–119.
- Soon CS, Brass M, Heinze HJ, Haynes J-D (2008) Unconscious determinants of free decisions in the human brain. *Nat Neurosci* 11:543–545.
- Stern ER, Gonzalez R, Welsh RC, Taylor SF (2010) Updating beliefs for a decision: neural correlates of uncertainty and underconfidence. *J Neurosci* 30:8032–8041.
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.
- Todd PM, Gigerenzer G (2007) Environments that make us smart: ecological rationality. *Curr Dir Psychol Sci* 16:167–171.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Usher M, McClelland JL (2001) The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev* 108:550–592.
- van Maanen L, Brown SD, Eichele T, Wagenmakers EJ, Ho T, Serences J, Forstmann BU (2011) Neural correlates of trial-to-trial fluctuations in response caution. *J Neurosci* 31:17488–17495.
- van Veen V, Krug MK, Carter CS (2008) The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *J Cogn Neurosci* 20:1952–1965.
- Wallis JD (2007) Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* 30:31–56.
- 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.
- Yang T, Shadlen MN (2007) Probabilistic reasoning by neurons. *Nature* 447:1075–1080.