

# Retrospective Valuation of Experienced Outcome Encoded in Distinct Reward Representations in the Anterior Insula and Amygdala

 Martin D. Vestergaard and  Wolfram Schultz

Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom

Our ability to evaluate an experience retrospectively is important because it allows us to summarize its total value, and this summary value can then later be used as a guide in deciding whether the experience merits repeating, or whether instead it should rather be avoided. However, when an experience unfolds over time, humans tend to assign disproportionate weight to the later part of the experience, and this can lead to poor choice in repeating, or avoiding experience. Using model-based computational analyses of fMRI recordings in 27 male volunteers, we show that the human brain encodes the summary value of an extended sequence of outcomes in two distinct reward representations. We find that the overall experienced value is encoded accurately in the amygdala, but its merit is excessively marked down by disincentive anterior insula activity if the sequence of experienced outcomes declines temporarily. Moreover, the statistical strength of this neural code can separate efficient decision-makers from suboptimal decision-makers. Optimal decision-makers encode overall value more strongly, and suboptimal decision-makers encode the disincentive markdown (DM) more strongly. The separate neural implementation of the two distinct reward representations confirms that suboptimal choice for temporarily extended outcomes can be the result of robust neural representation of a displeasing aspect of the experience such as temporary decline.

**Key words:** amygdala; anterior insula; decision utility; experienced utility; neuroeconomics; suboptimal choice

## Significance Statement

One of the numerous foibles that prompt us to make poor decisions is known as the “Banker’s fallacy,” the tendency to focus on short-term growth at the expense of long-term value. This effect leads to unwarranted preference for happy endings. Here, we show that the anterior insula in the human brain marks down the overall value of an experience as it unfolds over time if the experience entails a sequence of predominantly negative temporal contrasts. By contrast, the amygdala encodes overall value accurately. These results provide neural indices for the dichotomy of decision utility and experienced utility popularized as *Thinking fast and slow* by Daniel Kahneman.

## Introduction

When considering whether to revisit a previous holiday destination, economic theory holds that you compare it with other previous holidays and choose the destination that offered the best holiday in the past (Von Neumann and Morgenstern, 1947).

However, summarizing the overall value of an experience that unfolds over time is not trivial. Behavioral economics and social psychology studies have indicated that our impression of overall value is often dominated by the outcome in the final moments for both positive and negative experiences (Fredrickson and Kahneman, 1993; Redelmeier and Kahneman, 1996; Baumgartner et al., 1997; Ariely, 1998; Fredrickson, 2000; Schreiber and Kahneman, 2000; Do et al., 2008). A holiday with steadily improving weather may be experienced as more pleasant than one with declining weather, and you may therefore end up preferring the shorter holiday simply because of the distribution of sunny days. This effect presents a problem, not only for holidaymakers but also for everyone else deciding whether a particular previous experience of any sort merits repeating. While it is well known that people generally prefer increasing outcomes (Vestergaard and Schultz, 2015) and that perceived reward trends can be used to guide foraging decisions (Wittmann et al., 2016), no previous

Received Sep. 3, 2019; revised Aug. 12, 2020; accepted Aug. 25, 2020.

Author contributions: M.D.V. and W.S. designed research; M.D.V. performed research; M.D.V. analyzed data; M.D.V. and W.S. wrote the paper.

This work was supported by Wellcome Trust Grants 095495 and 204811 and the European Research Council Advanced Grant ERC-2011-AdG 293549. We thank Kelly Diederer for her generous help and advice and Marta Correia, Matt Davis, Paul Fletcher, Rik Henson, Russell Thompson, and Valerie Voon for support and advice with neuroimaging. We also thank the two anonymous referees who offered constructive commentary on an earlier version of the manuscript.

Requests for data and software should be directed to the corresponding author (M.D.V.).

The authors declare no competing financial interests.

Correspondence should be addressed to Martin D. Vestergaard at mdv23@cam.ac.uk.

<https://doi.org/10.1523/JNEUROSCI.2130-19.2020>

Copyright © 2020 the authors

study has examined how the human brain summarizes sequences of experienced rewards.

Inaccurate summary valuation leads to suboptimal choice for temporally extended outcomes. Poor decision-making is sometimes explained in terms of conflicting intuitions around intertemporal choice (Frederick and Loewenstein, 2008), suboptimal risk aversion (Kuhnen and Knutson, 2005), or competing concerns between short and long-term future goals (Hare et al., 2009). However, it remains unexplored how these ideas possibly relate to retrospective valuations. In a previous study, we have shown that the value experienced while an outcome is received is not the same as the incentive value later manifested in choices to consume the outcome again (Vestergaard and Schultz, 2015). The central idea is that experienced and incentive values embed independent impressions of an outcome's hedonic impact, such as its duration and temporal profile (Kahneman, 2003). Thus, in the example above, the impression of an outcome's temporal decline should be ignored when summarizing experienced value, and if it cannot be ignored, then it must be encoded in the brain. Hence, we sought to characterize how brain activity acts to regulate the distinct value signals for experienced outcomes and how these mechanisms can lead to suboptimal choice. Studies of the neurobiology of different value signals have often focused on the prefrontal cortex (PFC) and its subdivisions (Hare et al., 2008, 2009; Kuo et al., 2009; Donoso et al., 2014; Rich and Wallis, 2014), and since there are no strong predictions regarding the neurobiology of retrospective summary evaluations, we focus broadly on the PFC and limbic system. The ventral striatum (VStr) receives inputs from amygdala, PFC and insular cortex that share prominent roles in reward systems (Robbins and Everitt, 1996; Paton et al., 2006; Morrison and Salzman, 2010; Haber, 2011). We therefore hypothesize that activity in these brain structures assigns preferential significance to the summary value for outcomes that get better over time. We investigated the neurobiology underlying summary valuation of experienced outcome, and we examined how these processes differ with optimality of decision-making.

## Materials and Methods

### Subjects

Twenty-eight healthy male volunteers participated in the experiments. Male participants were used to minimize effects of cyclic modulation in risk attitude (Lazzaro et al., 2016). They were 21–36 years old (average 25.9, SD 3.7), with no history of neurologic or psychiatric disease, closed head injury, no self-reported substance abuse or use of psychoactive medication, and with normal or corrected-to-normal vision. One subject was excluded from the analyses because of excessive movement during scanning (>20 mm). The participants were recruited to take part in a gambling experiment and they were naive to the main purpose of the study. All participants provided written, informed consent. They were paid a fixed fee to participate (£5/h for the behavioral preexperiment and £10/h for the neuroimaging experiment) plus a variable amount of prize money (£5–15) according to task performance in each experiment (see Reinforcement schedule).

### Experimental rationale and protocol

The subjects attended 2 days of experimentation on a monetary valuation and choice task. We used a monetary incentive because it is known to engage the PFC and Striatum (O'Doherty et al., 2001; Elliott et al., 2003; Knutson et al., 2005). On the first day of experimentation, the participants received instructions about the tasks, performed the preexperimental valuation task (see below) and practiced the main experimental task (see below) in an IAC double-walled, sound-attenuated test booth. The preexperimental valuation lasted approximately 15 min in total.

Training on the main valuation task followed the protocol of experiment 2 in the previous study (Vestergaard and Schultz, 2015) and included all control and experimental conditions. Training on the main experimental task lasted approximately 2 h and was divided into 12-min blocks between which the participants were given an optional break. No feedback on performance in the main task was offered at this stage. Following successful completion of preexperimental valuation and training, the participants were booked into a scanning slot to take place no more than 7 d later. They then did the main experimental task in the scanner, this time divided into two scan sessions lasting approximately 30 min each. Between the two scan sessions was an intermission, in which the participants remained in the scanner and were allowed a pause to rest. After the scan sessions, they received feedback on their performance, and they received total payment in cash. The protocol was approved by the Cambridge Research Ethics Committee under reference number 04/Q108/190.

### Stimuli

In the main experimental task, associations were formed between visual conditioned stimuli (CSs) and unconditioned stimuli (USs). The CSs consisted of abstract figures composed by arranging randomly squares and triangles of four different colors of equidistant hue, each  $50 \times 50$  pixels. Thus, the CS was  $200 \times 200$  pixels. The USs consisted of sequences of gold coins presented at a stimulus onset asynchrony of 350 ms. Each coin was presented on a background with the same average color as the coin so that scaling of the coin did not result in variation in the average color spectrum of the stimuli. The coins varied in simulated volume. The temporal profiles were composed by scaling a sigmoid function to first generate a decreasing reference profile  $Q_n = F + (I - F) / [1 + \exp(s(n - N/2))]$  with  $N \in \{15:19\}$  elements, steepness  $s = 2$ , initial scale  $I = 0.7$ , and final scale  $F = 0.3$ . The experimental profiles were then composed by calculating obfuscated magnitudes,  $m_n = Q_n + O$  where the obfuscation noise  $O \in N(0, \sigma^2)$  was added to make less obvious the underlying temporal profile. Increasing profiles were generated by inverting the decreasing sequence along the temporal dimension and dominated alternatives were generated by removing elements from the longer profile. Following each sequence was a visual mask composed by scrambling the image of the reference coin on background. The association between CS and US was constant within one trial only (i.e., new CSs and USs were used on every trial). In the preexperimental valuation, single coins of varying size were presented for 350 ms.

### Preexperimental valuation

We measured the relationship between the physical size and experienced value of the virtual coins. The experienced value of a reward stimulus scaled in size to the magnitude  $m$  is given by Stevens' Power Law (Stevens, 1957):  $x = A m^\alpha$ , where  $A$  and  $\alpha$  are individual parameters of the observer. The objective value of a virtual coin is simulated volume ( $V = 100 m^3$ ), and the value function relating a pot of coins of scales  $m(t)$  to experienced value is therefore  $x(t) = K V^\kappa(t)$  where  $\kappa = \alpha/3$ ,  $K = A/(100^\alpha)$ . The coins were presented to the subjects in a 3D projection to simulate variation in volume (Fig. 1). A coin scaled in diameter by a factor  $m$  therefore differs in simulated volume by  $m^3$  compared with an unscaled coin, and insofar as the specific value of a precious coin is given in prize units per mass unit (e.g., £/g), the objective value is proportional to volume. Thus, if experienced value corresponded to objective value, there would be a linear relationship between simulated volume and subjective valuations (i.e.,  $\kappa = 1$ ). The experienced value of the virtual coins was determined by a Becker-DeGroot-Marschak (BDM) valuation task, which was designed to secure the participants an endowment of coins serving as gambling tokens in the main experimental task (see below) and which measured the participants' willingness to pay (WTP) for the coins.

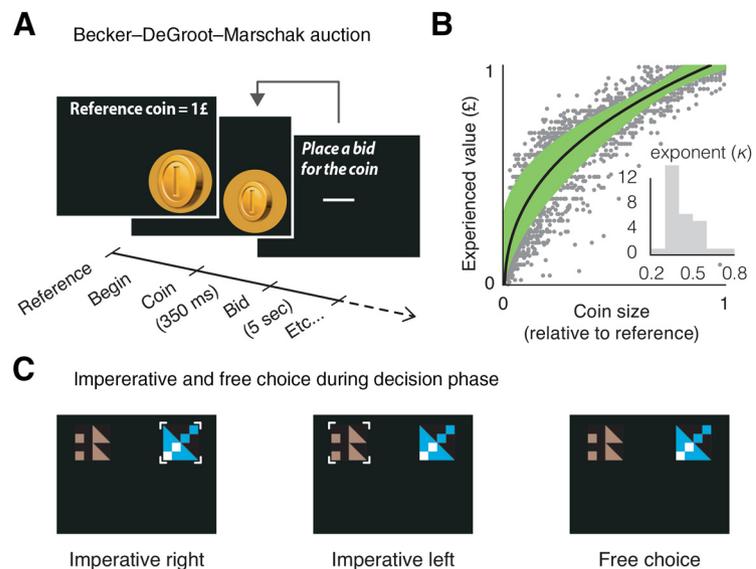
The participants were given a 5£ budget, and they were instructed to place bids on each of 120 coins according to how much they felt each coin was worth. The coins varied in simulated volume from 1% to 100% of a reference coin that was shown on the screen before the bidding started. After the bidding round, a randomly chosen subset of the bids were drawn to exchange the 5£ budget to approximately 30 virtual coins of varying size and value according to a second-prize auction (Vickrey,

1961). This set of coins would become the subject's endowment in the main experiment that followed. Before the bidding round, the reference coin of nominal value 100 pence (£1) was shown on the screen. A training round of 20 coins preceded the actual bidding round. Each coin was visible for 350 ms following a response window showing a black screen with the text "Place a bid for the coin" for up to 5 s (Fig. 1A).

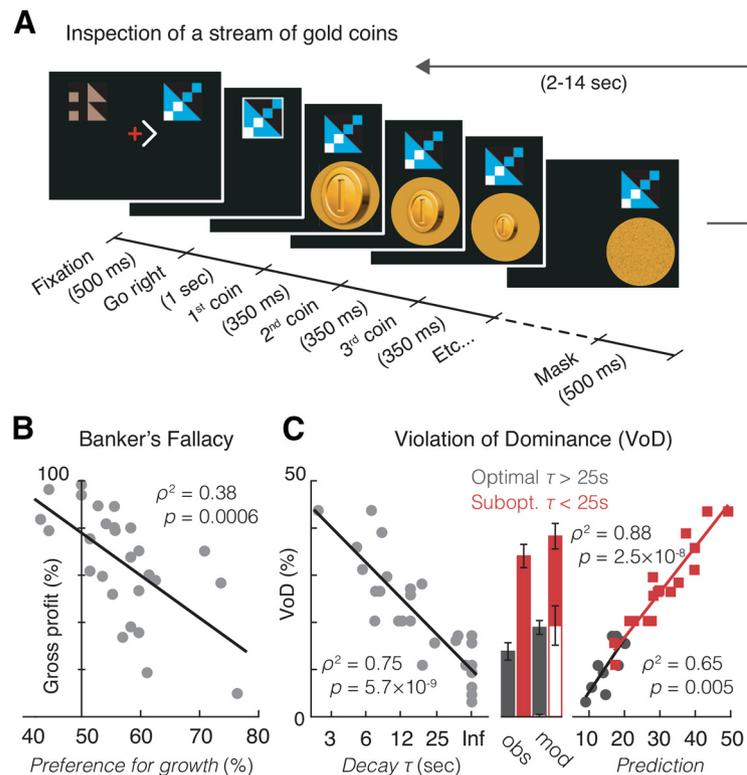
After the bidding round the participants watched a computer animation of the BDM auction, in which one coin at a time was drawn from the total pool and the associated bids placed by the participant were shown against the computer's bids. For each coin drawn in this way, the computer placed a random bid drawn from a rectangular distribution without taking into account the specification of the coin at stake or the participant's bid. If the computer's bid was higher, the coin at stake was discarded and a new coin was drawn from the pool of bids without replacement. If the participant's bid was higher, they would buy the coin and they would pay the amount bid by the computer. This amount was then taken from their budget, and this procedure would continue until the budget was spent. In this way approximately 30 bids were realized to ensure the participants an initial endowment of virtual coins. This endowment had an expected value of £10 because the coins were obtained in a second-price auction (Vickrey, 1961) consistent with the BDM method (Becker et al., 1964).

#### Main experimental task

In the main experiment, we used a monetary venture with explicit choice (Vestergaard and Schultz, 2015). On each trial, the participants were offered the choice of one of two competing options indicated by two CSs. They were instructed to first inspect one of the options shown by a white arrow. Then followed a sequence of coins of varying sizes (Fig. 2A). After inspection of each pair of options, the two CSs were shown again and the participants indicated which sequence of coins they preferred in a free choice and in two imperative choices presented in random order (Fig. 1C). The free choice served to record their revealed preference whereas the imperative choices allowed us to analyze Pavlovian associations relating to both options. It was explained to the participants to approach the task in the following way: "two pots of money are on offer; first you must inspect the contents of each pot and then choose one or the other." Each pair of options consisted of a sequence of gold coins and a subset from that sequence presented either decreasing or increasing. On each trial a decreasing sequence was in competition with an increasing sequence and either could be long while the other was short. Thus, the options differed quantitatively by the value of the coins omitted from the longer sequence to produce the dominated alternative and qualitatively in the order in which the coins were presented. Between trials, the long sequences varied in length between 15 and 19 coins,



**Figure 1.** Preexperimental valuation and decision phase of the main experiment. **A**, The experienced value of the virtual coins was obtained by measuring the participants' WTP in a BDM auction. **B**, Individual BDM bids (gray dots), mean value function (black line), 95% confidence interval (green area), and distribution of estimated value function exponents,  $\kappa$  (inset). **C**, Imperative and free choice screens in the main experiment.



**Figure 2.** Inspection task and behavioral results. **A**, Two alternative coin streams are inspected sequentially. **B**, The Banker's fallacy is the tendency to prefer the growing option. Since the expected value is the same for growing and declining options, preference for growth leads to diminished profit. Individual mean data (gray dots). **C**, Prevalence of violation of dominance (VoD) and its relation to the time constant of the leaky integrator (left); observed and predicted VoD (%) for optimal ( $\tau > 25$  s, gray) and suboptimal ( $\tau < 25$  s, red) decision-makers (middle). The bars show VoD including softmax errors; the unfilled part shows cases in which the model predicts VoD because of leaky integration. Correlation between predicted and observed VoD for optimal and suboptimal choosers (right).

and to each long sequence, a weakly dominated alternative was created by removing between zero and four coins creating short sequences of 11–15 coins. Thus, the weakly dominating option was always at least as good as the alternative (Nurmi, 2006).

Two control conditions were used during the preexperimental training: (1) a decreasing versus an increasing sequence with the same number of coins (11–19 coins) and (2) a short (11–15 coins) flat sequence versus a longer (15–19 coins) flat sequence ( $s=0$ ). The data from the control conditions were included in the model fitting of the parameters,  $\tau$  and  $B$  in the decision model below. As previously, we observed strong preference for the longer and for the increasing univariate control sequences: for the control sequences of same duration the average preference ( $\pm$ SEM) for the increasing rewards was 0.63 ( $\pm$ 0.04), and for the flat sequences, the average preference for the longer reward was 0.81 ( $\pm$ 0.03). The duration of the inspection epochs varied from 3.85 s (11 coins) to 6.65 s (19 coins).

To obtain insight into the participants' final understanding of the structure of the task, we interviewed them after they had received final payment at the end of the experiments and asked how they had approached the task. Most participants reported a tendency to rely on a strategy whereby coins were classified in two or three bins (e.g., small/medium/large) where the number of large coins became the main determinant of choice. Some participants also reported that sometimes the large coins seemed to come early and other times later, and they were mindful not to let that affect their judgment. No participant reported relying entirely on duration, option order or screen position.

#### Reinforcement schedule

The total value of a pot of money was the sum of the experienced values of each coin in the coin stream. One of the streams comprised coins taken from the participant's endowment while the other was on offer from the bank. The preferred pot of money would go back into the endowment but the coin stream was not shown again. In this way, the value of the endowment could either increase or decrease by the difference in total value between the two options, or the participants could break even depending on their choice and the respective funding sources for the two options, which they did not know of. At the end of the experiment, payment was calculated by realizing four randomly chosen trials. The value of the initial endowment was then adjusted according to the participants' performance in the four randomly chosen trials, and payment was made on the basis of the adjusted value of the endowment at the end of the experiment. In the main experimental task described above, the participants all experienced the same amount of identical coins, but their goal and incentive values were calculated according to individual parameters as detailed in the decision model below.

#### Decision model

##### Value signals

Two prominent value signals that may guide decision-making are goal value (GV) and decision value. While GV represents the benefit of an outcome in relation to motivational state, decision value relates this benefit to the cost or effort involved in acquiring it (Hare et al., 2008). In our study, the sum of the experienced values in a coin stream is the GV (Eq. 1), but there is no cost or effort associated with acquiring a coin stream. Instead there is a disincentive manifested in people's distaste for declining sequences. Below we define the incentive value (IV) (Eq. 3) as the value revealed by participants to have motivated their choices, and the disincentive markdown (DM; Eq. 4) as the penalty imposed by participants on a coin stream when they reveal their preference. The decision model is identical to the one derived in a previous study (Vestergaard and Schultz, 2015). Thus, the total GV of a coin stream is:

$$GV = \int x(t) dt. \quad (1)$$

The IV of a stream of gold coins is continuously tracked in relation to historical incentive. The rate of change in IV,  $dy/dt$  is the experienced value  $x(t)$  marked down in relation to previously accumulated incentive:

$$dy/dt = x(t) - wy(t), \quad (2)$$

where  $w$  is the immediacy of the markdown on the experienced values. Thus, the total IV of a coin stream of duration  $T$  is obtained by leaky integration of the experienced values:

$$IV = \int_0^T x(t) e^{w(t-T)} dt, \quad (3)$$

where  $\tau = 1/w$  is the decay constant of a leaky integrator. In other words, the salience function (Tsetsos et al., 2012) of the sequence is an exponential filter. GV and IV differ by the total DM:

$$DM = GV - IV, \quad (4)$$

which represents the quantification of a penalty imposed on a coin sequence depending on its temporal configuration. The effect can also be regarded as a contrast effect, and Equation 3 may therefore be called contrast-guided retrospective valuation. Preference  $P$  for one of two competing options (a, b) is given by logistic discrimination:

$$\log\left(\frac{P}{1-P}\right) = B(\epsilon + \beta), \quad (5)$$

where  $\epsilon = -\log(IV_b/IV_a)$  is the incentive evidence in favor of option a,  $B$  is the inverse temperature of the decision process, and  $\beta$  is its bias. In discrete notation, IV can also be expressed in its recursive form:

$$y_n = (1-a)y_{n-1} + x_n, \quad (6)$$

where  $a = 1 - \exp(-w)$  as illustrated in Figure 3A.

The leaky integrator has been used as generative model for many different phenomena in a wide range of disciplines within engineering, psychology, and physiology (Hodgkin and Huxley, 1952). It imposes a greater markdown on long sequences than on short sequences, so for shorter sequences to become preferable their early elements must be small.

In a previous study, we considered many alternative models allowing for bias for increasing options and option order, as well as combinations of biases with or without leaky integration as described above, and we discussed biological aspects of their implementation in humans and other animals (Vestergaard and Schultz, 2015). The result of model comparison in the previous study was that bias-free leaky integration was the most likely mechanism to describe suboptimal preference for temporally extended outcomes. Specifically, there was no effect of option order indicating any systematic memory effect. Consequently, in the current implementation  $\beta = 0$ . A functional interpretation of these equations is that the markdown (DM) describes the varying disincentive effect of the temporal configuration of an extended outcome. A long sequence will incur a greater absolute markdown than a short sequence, and an increasing sequence will incur a smaller markdown than a flat sequence that again will incur a smaller markdown than a decreasing sequence. It is this differential markdown that can explain preference for a shorter increasing sequence over a longer decreasing sequence. Thus, neither the behavioral results nor the decision model distinguishes whether the differential markdown reflects a subjective aversion to decline or a craving for growth.

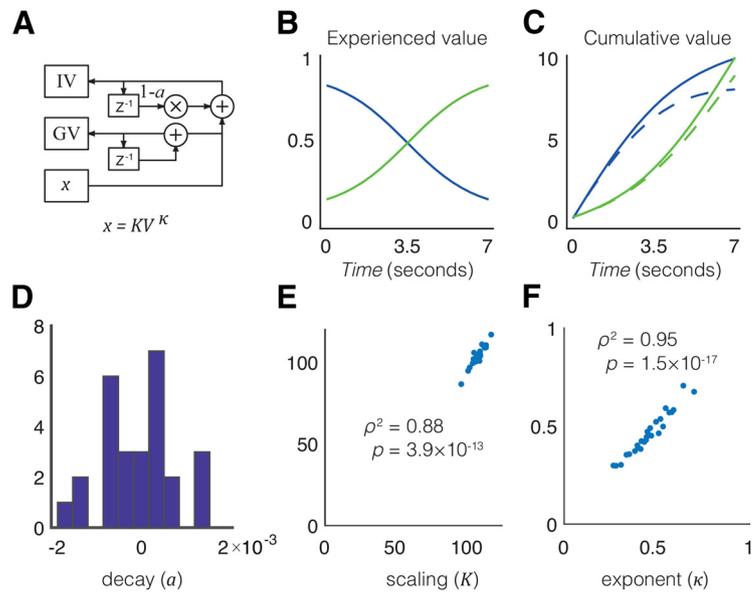
Compared with bias models or models that involve a different generative mechanism for increasing and decreasing sequences, our model (Eq. 2) operates on any sequence, and it is not necessary for the decision-maker to keep track of whether a sequence is increasing or declining. This parsimony supports our model's plausibility over more complicated mechanisms. Thus, contrast-guided retrospective valuation as described above remains a biologically plausible mechanism for constructing summary valuation of temporally extended outcomes. Using this model, individual performance in the task is chiefly characterized by two parameters:  $\kappa$  that determines how compressive is the individual value function, and  $\tau$  that characterizes how leaky is the integration of experienced value. The lower is the value of  $\tau$ , the more are the early instances of a temporally extended outcome marked down, and thus the higher is the risk of committing to the choice of a dominated option. We can therefore regard  $\tau$  as a marker across a continuum of optimality with low values associated with a high degree of suboptimality. In this report, we use the term suboptimal for choices that violate the dominance axiom of economic theory (Barbera et al., 2004).

We can score performance according to whether observed choices optimize GV so that any choice that does not optimize GV is regarded as an error. Decision errors happen either because of decision noise in cases where the values under consideration are very similar (softmax errors) or as a result of the leaky integration yielding evidence in favor of an inferior option. Thus, the IV model can be used to predict decision errors, and it can quantify the proportion of decision errors that are simply softmax errors (Eq. 5). Systematic decision errors, over and above softmax errors, represent a violation of dominance because the rejected options are at least as valuable as the preferred options. Equations 1–4 can thus be used to characterize cases of violation of dominance to understand what conditions lead to suboptimal choices.

#### Individual value functions

To quantify the experienced value of the gambling tokens, we fitted power value functions to the data from the BDM auction (Fig. 1A). The WTP data themselves are noisy (Fig. 1B) so the power function is used to obtain a monotonous increasing non-saturated value function. The individual value functions were all concave within a wide range of exponents ( $0.298 < \kappa < 0.759$ , average 0.438, SD 0.109; Fig. 1B). Using these value functions, we calculated the total GV of each pot of money inspected in the gambling task (Fig. 2A). We then fitted the IV model to the choice data. The GV model is simply the sum of the experienced values for each coin. By contrast, the IV model marks down the earlier coins in a sequence leading to relative overvaluation of growing coin streams compared with declining coin streams. The extent of this markdown is determined by the time constant ( $\tau$ ) of the leaky integrator. Using the GV model, we further calculated the gross profit obtained by each participant relative to the maximum attainable. A disproportionate focus on growth is a “Banker’s fallacy” when tolerance to experienced decline would result in a higher profit (Fig. 2B). We characterized the severity of individual behavior as the proportion of violation of dominance in the choice data. We found a high correlation between violation of dominance and gross profit ( $\rho^2 = 0.96$ ,  $p = 2.9 \times 10^{-19}$ ) and between Banker’s fallacy and violation of dominance ( $\rho^2 = 0.38$ ,  $p = 0.0006$ ). Moreover, between participants, the average loss incurred in dominated choices was positively correlated with the degree of violation of dominance ( $\rho^2 = 0.54$ ,  $p = 0.000011$ ). In other words, those who strongly preferred growth made more bad choices, and the average loss in their bad choice was greater than in those who did not strongly prefer growth but still made occasional bad choices.

As predicted, there was a systematic relationship between the decay constant ( $\tau$ ) in the IV model and violation of dominance ( $\rho^2 = 0.75$ ,  $p = 5.7 \times 10^{-9}$ ). In the more optimal decision-makers ( $\tau > 25$  s), the gross profit was close to 100% (Fig. 2B) as in these cases the incidence of violation of dominance was low. In the more suboptimal decision-makers, there was large variation in decay ( $2.3 \text{ s} < \tau < 25 \text{ s}$ ) and higher incidence of violation of dominance (Fig. 2C, left). More optimal decision-makers showed an average ( $\pm$ SEM) error rate of 0.11 ( $\pm 0.017$ ), whereas the more suboptimal showed an error rate of 0.27 ( $\pm 0.023$ ). The IV model predicted total error rates of 0.15 ( $\pm 0.012$ ) in optimal and 0.31 ( $\pm 0.02$ ) in suboptimal decision-makers of which 0.002 ( $\pm 0.002$ ) and 0.16 ( $\pm 0.034$ ), respectively, were predicted violations of dominance (Fig. 2C, middle). While the model slightly overpredicts error rates, the correlation between observed and predicted error is high and no different between the optimal and suboptimal ( $\chi^2 = 1.84$ ,  $p = 0.088$ ; Fig. 2C, right). These results support the notion that violation of dominance in the optimal decision-makers is not a result of Banker’s fallacy but simply perceptual errors in cases where the GV of the two options were very similar.



**Figure 3.** Dissociation between IV, GV, and experienced value ( $x$ ), and the potential effects of decay on IV and  $x$ . **A–C**, GV and IV by leaky and non-leaky integration of experienced value ( $x$ ). **A**, Schematic network of the recursive formulation of IV in Equation 6. **B**, Two experienced options with equal contents, one stream follows a growing trend (green) while the other declines (blue) over time. **C**, Cumulative GV (solid) and IV (dotted). Differential leak leads to IV in favor of the increasing option. **D–F**, Effect of integrating decay in valuation of single coins. **D**, Distribution of decay estimates ( $a$ ). **E**, **F**, Relationship with non-leaky parameter estimates,  $\kappa$  and  $K$ .

The mechanism described above assumes that leaky integration is a feature of sequence evaluation and that the power value function operates separately on the individual coins regardless of their position in a sequence. To analyze whether perception of the size of individual coins was nonetheless affected by past events, we combined the power value function with the recursive form of IV (Eq. 6),  $y_n - y_{n-1} = K V_n^\kappa - a y_{n-1}$ , where the marginal incentive,  $y_n - y_{n-1}$ , in this case is taken as the WTP for the individual coin of size  $V_n$ . We then refitted the  $K$  and  $\kappa$  parameters of the power value function together with the recursive decay  $a$  to the auction data. If perception of individual coin size is affected by the historical sequence, this effect can be measured by the recursive decay parameter,  $a$ , in exactly the same way as  $\tau$  measures the effect of leaky integration, as illustrated in Figure 3A. The distribution of recursive decay parameter estimates obtained in this way and a comparison between  $K$  and  $\kappa$  parameter estimates obtained with and without the recursive value function are shown in Figure 3D–F. This analysis resulted in negligible parameter estimates for the recursive decay  $-1.9 \times 10^{-3} < a < 1.5 \times 10^{-3}$ , and convincingly reproduced the parameter estimates for the power value function ( $K$ :  $\rho^2 = 0.88$ ,  $p = 3.9 \times 10^{-13}$  and  $\kappa$ :  $\rho^2 = 0.95$ ,  $p = 1.5 \times 10^{-17}$ ). These results show that when the participants observed streams of coins in this way, there was no effect of past events. Thus, the decay parameter cannot account for perception of the size of the single coins, and it therefore seems unlikely that behavioral and neural effects of  $\tau$  depend exclusively or critically on perceptual function.

#### MRI data acquisition and preprocessing

MRI data were acquired at the MRC Cognition and Brain Sciences Unit (CBU) in Cambridge, UK, on a Siemens Trio Tim 3-Tesla scanner using a 32-channel head coil. An MPRAGE sequence was used to acquire a whole-brain T1-weighted structural image (TR = 2.25 s, TE = 2.98 ms, flip angle 9°, 192 slices,  $1 \times 1 \times 1 \text{ mm}^3$ , FOV  $256 \times 256 \text{ mm}^2$ ), and functional data were acquired with an echoplanar imaging (EPI) sequence (TR = 2.03 s, TE = 30 ms, flip angle = 78°, 33 axial slices of matrix  $64 \times 64$ , in-plane resolution  $3 \times 3 \text{ mm}^2$ , thickness 2 mm, gap 1 mm, FOV  $192 \times 192 \text{ mm}^2$ ). To optimize sensitivity in the orbital frontal cortex (OFC), we used a tilted acquisition of 30° relative to the anterior-posterior commissures line (Deichmann et al., 2003). T2\*-weighted EPIs

were acquired over two sessions, resulting in up to 1070 volumes per session depending on the duration of the self-paced experimental task. Field maps (TR = 400 ms, TE = 5.19/7.65 ms, flip angle 60°, 3 × 3 × 3 mm<sup>3</sup>, FOV 205 × 205 mm<sup>2</sup>) were acquired in the pause between the two EPI sessions.

Data were analyzed with SPM12 (Wellcome Department of Imaging Neuroscience, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>) using AA 4.2 (automatic analyses; Cusack et al., 2014). Preprocessing with a standard MRC-CBU recipe (Taylor et al., 2017) included image unwarp using individual field-maps, realignment, slice-time correction, co-registration of functional images to the T1-weighted structural image, unified segmentation (Ashburner and Friston, 2005), DARTEL normalization (Ashburner, 2007) to the Montreal Neurologic Institute (MNI) template and spatial smoothing using a Gaussian kernel, 8 mm wide at half maximum. The time-series from each scan session were high-pass filtered (1/128 Hz) and serial autocorrelations were estimated using an AR(1) model.

### fMRI general linear model (GLM)

We used a single GLM to analyze blood oxygenation level-dependent (BOLD) activity measured during the inspection and choice phases (Table 1). The GLM used a 2 × 2 × 2 factorial boxcar specification of the inspection epochs. The factors were: option order (first/second), valence (growth/decline), and choice (prefer/reject). Preference was assigned to each inspection sequence based on the choice revealed in the subsequent decision phase of each trial.

Delta events were used to indicate the end of the inspection epochs (Fig. 2A) and to indicate the free and imperative choices of the decision phase (Fig. 1C). The two imperative choices were encoded as “preferred” and “rejected” based on the preference revealed in the free choice.

All regressors were included in the single GLM with parametric modulators for the different types of values (see below) added to the delta events. Motion parameters from the realignment preprocessing step, response times and trial number were used as covariates of no interest, and separate intercepts were estimated for each of the two scanning sessions.

### fMRI contrasts

The two inspection epochs differ qualitatively in that a relative valuation of an option’s content was only possible during inspection of the second option. We refer to this difference between the first and second inspection as “naive” or “comparative.” We use the direct contrast of the two to first ascertain that the inspection task engages known reward circuitry regardless of whether the inspected coin streams were increasing or decreasing. To obtain an anatomically meaningful separation of the activation cluster we used the most conservative family-wise error (FWE) correction at the voxel level (Extended Data Table 5-1). We then asked whether the predefined VStr regions of interest (ROIs; see below) were engaged in encoding preference during inspection, and whether any region was preferentially engaged for growing sequences (Extended Data Table 5-2) and whether any region encoded interaction between preference and growth (Extended Data Table 6-1). These analyses were intended to show the neural underpinnings of valuation during inspection and their specific involvement in encoding preference and option valence.

We then analyzed the earliest point in time when the total summary might be encoded. We have assumed that summary value can be constructed on the fly (Eq. 2); thus, the end of the cumulative process is the earliest point for the total summary, and we report the encoding of the comparative summary (Extended Data Table 7-1) and the parametric modulation of GV and DM on the end of inspection (Extended Data Table 7-2).

To analyze overt choice, divorced from any action leading to its execution, we used the contrast between the free choice (indicating preference) and the imperative choice action for the preferred option. Similarly, to analyze associative aspects of preference, regardless of choice, we used the contrast between the imperative choice action for the preferred option and the imperative choice action for the rejected option (Extended Data Table 8-1). For the connectivity analyses detailed

**Table 1. GLM events and regressors**

Event	Indicator	Parametric modulator	Design	Types
Inspection	Boxcar (4–7 s)		2*2*2	First, second Up, down Preferred, rejected
Summary	Delta	GV, goal value DM, disincentive markdown	2*2*2	First, second Up, down Preferred, rejected
Decision	Delta	dGV, goal value difference dDM, markdown difference	3	Free choice Imperative (preferred) Imperative (rejected)

In each trial, two options are inspected sequentially (first and second), one option is increasing (up) and the other decreasing (down), and one is subsequently chosen (preferred) while the other is not (rejected). The summary is defined as the end of the inspection epoch. In the decision phase, one choice is free, while two are imperative encoded according to the free choice.

below, we used the contrasts between the factorial levels of the inspection to analyze the effect of choice and valence on neuronal co-activation (Extended Data Tables 7-3, 7-4).

As mentioned above, all of the value regressors were included in the single GLM, and the GLM contrasts for inspection and choice were performed for these regressors (Table 1).

### Connectivity

We report the results of two connectivity analyses using psychophysiological interaction (PPI) during the inspection epochs. The central question to investigate is which areas in the brain are the sources of the differential markdown for increasing and declining sequences that can lead to suboptimal choice. The PPI analyses thus use the full factorial design of each inspection epoch assessing effects of the participants’ ultimate qualitative judgment (prefer/reject) in relation to sequence valence (growth/decline) and their interaction. The seed regions were anterior insula and amygdala activation identified for the summary evaluation (Extended Data Table 7-2). Eigenvariates ( $r = 12$  mm) were extracted and the interaction between brain activation and the three design contrasts (prefer/reject, growth/decline, prefer/reject\*growth/decline; Extended Data Tables 7-3, 7-4) on the whole brain was estimated using the PPI machinery implemented in SPM. We report random effects for the PPI analyses on the whole brain. Moreover, we display effect size by re-estimating PPI objects separately for each condition and extracting average slope estimates in 12-mm spheres around the coordinates identified in the random effects analyses for each PPI.

### Statistical analyses

#### Parametric modulators

GV and IV both vary as a function of the duration of a coin sequence. This collinearity means that GV and IV are correlated (mean  $\langle \rho^2 \rangle = 0.83$ ), and it would be difficult to interpret their differential leverage if they were both used as parametric modulators to analyze the functional MRI data. To address this problem, we characterized the neural aspect of the IV model by the difference between GV and IV that is the DM (DM = GV – IV). The main idea of this operation is to isolate the common variance because of the duration of a sequence in GV in order that DM may explain effects in the BOLD signal that do not depend on duration. Note that the shared variance stems from the effect of sequence duration that is directly related to the total value. The shared variance therefore has to be assigned to GV, because that way DM reflects only the extent to which the sequence declines regardless of how long it is. As mentioned above, the markdown can be thought of as the chooser’s dislike of certain temporal profiles, and this dislike is then imposed as penalty on the incentive compatible GV. To remove residual correlation between DM and GV, we included the Gramm–Schmidt orthogonalization implemented natively in SPM. This configuration of the parametric modulators does not alter the interpretation of effects of DM, and it ensures that GV remains unadjusted for effects of duration. This arrangement, which is the recommended use of orthogonalization (Mumford et al., 2015), also goes along with a semantic interpretation of

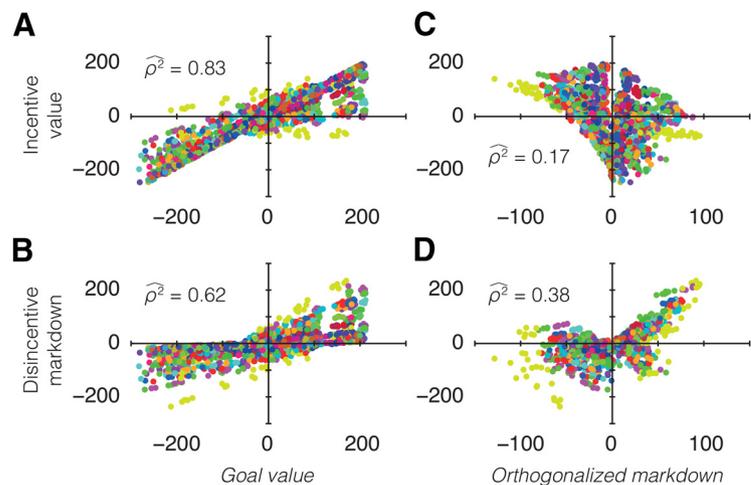
the sort of preference pattern that we would like to explain, namely, the “duration neglect” (Fredrickson and Kahneman, 1993).

The orthogonalized markdown remains positively correlated with the (un-orthogonalized) difference between GV and IV (mean  $\langle \rho^2 \rangle = 0.38$ ), and it is negatively correlated with IV (mean  $\langle \rho^2 \rangle = 0.17$ ; Fig. 4). We regressed the BOLD signal at the end-of-inspection events on GV and DM calculated in each participant (using individual estimates of  $\kappa$  and  $\tau$  as explained above) and for each inspected option. After the second inspection, value signals relating to both options are available so in the decision phase, we regressed the BOLD signal on the absolute differences in GV and DM between the two competing options.

The variation in  $\tau$  and  $\kappa$  between participants causes variation in the scale of the regressors, and SPM only uses mean centering to normalize parametric modulators. Although scaling a regressor does not change its predictive leverage at the first level of analysis, at the second level, variation in slope estimates might be confounded with variation in  $\tau$  and  $\kappa$  that would render between-subject analyses circular unless the parametric modulator is normalized at the first level. To address this issue, we  $z$  score normalized the parametric modulators in the first-level analyses.

#### Group analyses

To investigate variation in performance across participants, we classified the participants as either mainly optimal or mainly suboptimal decision-makers. As mentioned above, we use the term suboptimal for systematic decision errors that violate the dominance axiom of economic theory (Barbera et al., 2004). The IV model has the capacity to capture variance in the choice data reflecting violation of dominance insofar as its basis is disproportionate preference for increasing coin sequences. Therefore, we calculated the maximum decay value that would result in a difference in IV between increasing and decreasing coin streams of the same duration of at least 5%. This cutoff was chosen arbitrarily and resulted in a cutoff for  $\tau$  at 25 s. Therefore, we characterized the participants as either mainly optimal decision-makers ( $\tau > 25$  s,  $N = 10$ ) or mainly suboptimal decision-makers ( $\tau < 25$  s ( $N = 17$ ) based on individual decay time-constant estimates in the IV model. Seven out of the 10 participants in the optimal group identified in this way had an infinite decay value ( $\tau = \infty$ ). In the group analyses, we therefore included a binary covariate indicating the behavioral stratification ( $\tau < > 25$  s). This means that our main results are controlled for variation in  $\tau$ , and slope estimates of the covariate itself indicate the effects of optimality. Results obtained in this way were robust to alternative group definitions, such as a median split or a correlation in which participants with infinite decay values were given an arbitrary high value,  $\tau = 50$ . To analyze the extent to which effects identified in this way co-varied with  $\tau$  within the full range of estimated values, we conducted the following analyses: First, we used a leave-one-out method to estimate the peak coordinates in each participant by leaving out his data in a group analysis. Then, we extracted the average slope estimates in each participant in 12-mm spheres around the leave-one-out coordinates and calculated the correlation between average effect size and  $\tau < \infty$ . These analyses were intended to expose the extent to which interaction identified in the group analyses also more generally represented varying degrees of optimality within the entire operational range of the leaky integrator. The leave-one-out method for calculation of effect sizes aimed to effectively avoid circular analyses via cross-validation of activation coordinates (Kriegeskorte et al., 2009). For display purposes, we also use the stratification of participants as mainly optimal or suboptimal decision-makers to illustrate the direction of effects



**Figure 4.** Relationships between regressors (mean-centered values) and effects of orthogonalization. Each color shows data from one participant, and the mean square correlation coefficient ( $\rho^2$ ) is shown in each panel. **A, B**, Relation between GV and IV and between GV and markdown ( $DM = GV - IV$ ). Co-variation with sequence duration (number of coins) causes positive correlation between them. **C, D**, Relation between DM orthogonalized (with respect to GV) and IV, and between orthogonalized and unorthogonalized DM. When the correlation between DM and GV is removed, the orthogonalized DM is negatively correlated with IV and remains positively correlated with the unorthogonalized DM. Note that in participants with low markdown (i.e., high  $\tau$  values), DM contains little information after the common effect of duration has been removed, whereas in those with high markdown (low  $\tau$  values), there is a more linear relationship between the orthogonalized and unorthogonalized markdown.

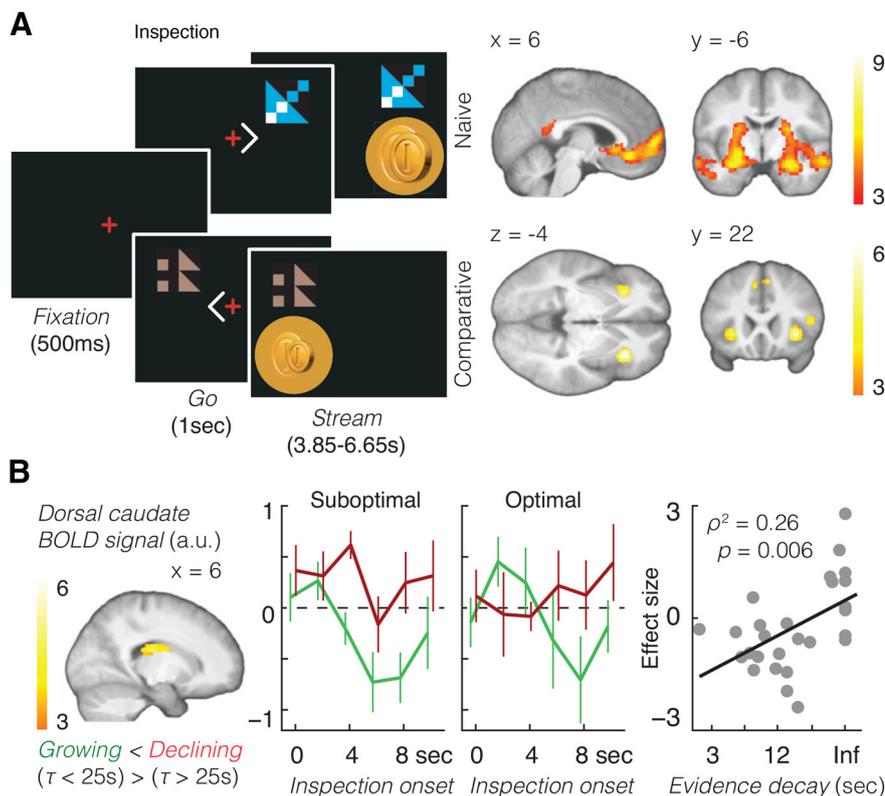
interacting with the decay constant, but we did not derive group statistics based on the subgroups.

We report random effects at  $p < 0.05$  that survive whole brain correction for FWE at the voxel level. Moreover, we estimated the minimum cluster size for FWE whole brain correction at the cluster level (Slotnick et al., 2003). We used a primary cluster-defining threshold of  $p < 0.001$ , which in some cases was increased to  $p < 0.0001$  or  $p < 0.00001$  to increase the spatial specificity of anatomically ambiguous activation. Although cluster-defining thresholds higher than 0.001 avoid the weakness in random-field theory that spatial autocorrelations are non-Gaussian (Eklund et al., 2016), we nevertheless adopted the conservative approach to increase the FWE threshold to  $p < 0.01$  for analyses using the most liberal cluster-defining threshold  $p < 0.001$ . This procedure gave cluster thresholds ranging from  $k > 68$  to  $k > 38$  ensuring high spatial specificity and low type I error rates (Woo et al., 2014). We also defined two ROIs in the VStr for which we used small volume correction ( $p < 0.05$ ) for a sphere of 16-mm radius. The ROIs were based on accumbens coordinates reported in a meta-analysis on the involvement of VStr in reward-related decision-making (Liu et al., 2011). We used the average coordinates of accumbens activation for evaluation in their tables 3 and 8, MNI [ $xyz$ ], RH: [12 8 -7], LH: [-14 7 -9]. These ROIs covered nucleus accumbens (NAc) including surrounding anteroventral putamen and caudate nucleus and extend into the ventral pallidum. To illustrate time courses of activation during inspection, a FIR model was used to estimate effect sizes in 12-mm spheres every TR after inspection onset, and unless otherwise indicated we show activation of  $t$  values at  $p < 0.001$  (unc.) for display purposes.

In the figures, activation maps were overlaid onto a structural image composed by stripping off the skull of the normalized individual structural scans in FSL and calculating the mean structural image of the participating subjects.

## Results

We used fMRI to study neuronal mechanisms in twenty-seven human volunteers engaged in a monetary valuation and choice task. The participants inspected two alternative pots of money each consisting of a flow of gold coins presented one coin at a time (Fig. 2A). Each pair contained a growing and a declining



**Figure 5.** Inspection of coin streams. **A**, Naive inspection involves bilateral activation in putamen, amygdala and ventromedial PFC, whereas comparative inspection specifically is correlated with anterior insula activation (Extended Data Table 5-1). **B**, Inspection of growth differentially activates dCN (Extended Data Table 5-2). Time-resolved effects are shown for illustration purposes. More suboptimal decision-makers show a downregulation in dCN for increasing coin streams compared with more optimal decision-makers, and the differential dCN activation correlates with the decay ( $\tau$ ) in individual incentive-value models.

stream of coins; one weakly dominated by the other. After inspection, the participants indicated which pot of money they preferred. The total value of any pot was independent of whether the coin stream was growing or declining, so optimal performance was achieved by disregarding the order in which the coins were experienced. Choosing a shorter stream of coins would be a violation of dominance, which we regard as an error. For a complete description of the task, see Materials and Methods.

### Behavioral support of a Banker's fallacy

In a behavioral preexperiment, the participants first evaluated the virtual gold coins in a second-price auction (Becker et al., 1964), and we calculated the GV of a pot of money as the sum of the individual coin values (Fig. 1). Based on preference data for coin streams presented along various temporal profiles, we estimated the parameters of a leaky-integrator that computed the IV of each pot of money. The IV predicts an individual's preference for increasing options (Vestergaard and Schultz, 2015). The difference between the GV and IV is a markdown that acts as disincentive on declining coin streams. The shorter the decay constant ( $\tau$ ) of the leaky integrator, the more pronounced is an individual's aversion to decline and the higher is therefore the error rate. We can therefore regard  $\tau$  as a marker across a continuum of optimality with lower values associated with a higher error rate. When the integrator leaks very slowly ( $\tau > 25$  s in these experiments), the GV and IV are very close to identical, indicative of near-optimal behavior.

Systematic preference for growth is a Banker's fallacy that leads to diminished profit (Fig. 2B). Decision errors occur either

because of decision noise, also known as "late noise" (Tsetsos et al., 2016), or because of an inclination toward growing coin streams of lesser total value. Results show that decision errors in the more optimal decision-makers are not a result of the Banker's fallacy but rather perceptual error occurring as a result of decision noise (Fig. 2C, middle). The leaky integrator predicts the Banker's fallacy, and individual  $\tau$  estimates are correlated with the error rate (Fig. 2C, left). A functional interpretation of these results is that temporally extended outcomes are experienced along two distinct reward representations: (1) the GV that is incentive-compatible, and (2) a competing DM on temporary decline.

### Neural implementations

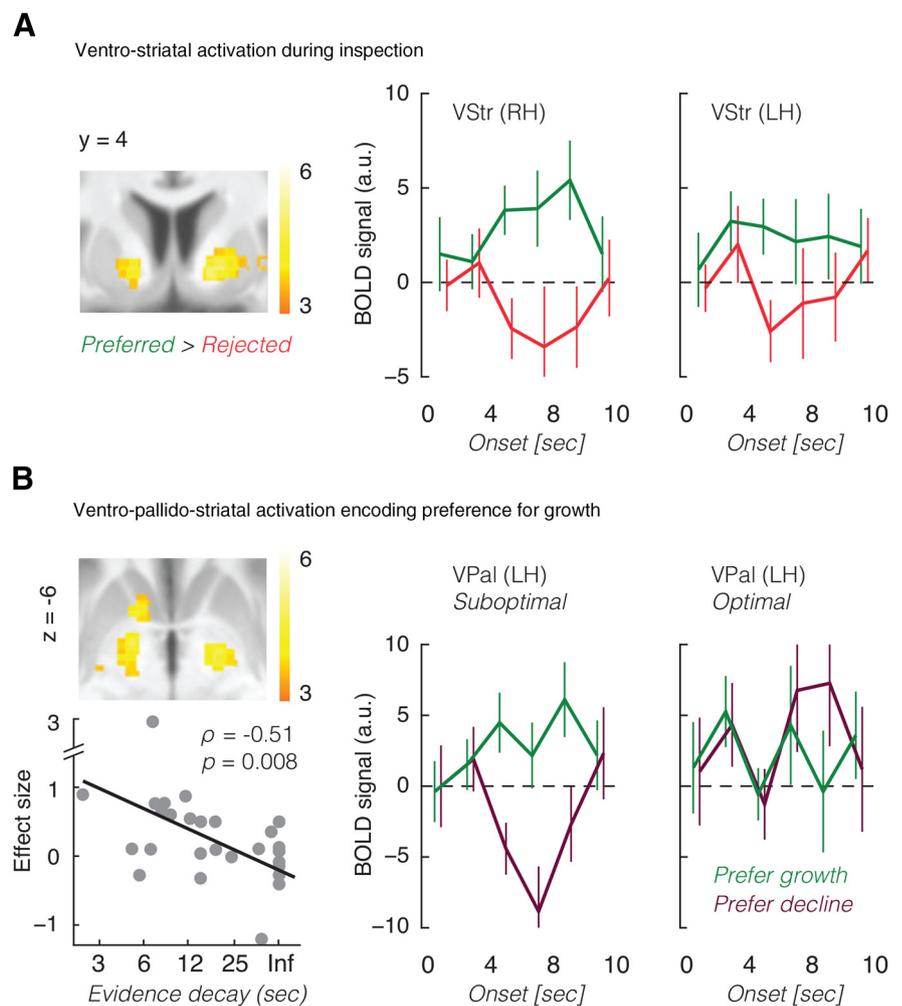
We investigated whether brain activation supported a neural implementation of a choice model with two distinct reward representations. We used the most conservative FWE correction at the voxel level as well as whole-brain correction at the cluster level (Slotnick et al., 2003). We first show that the task activates the reward system. Inspection of a sequence will engage vision and attention as well as the reward system, and to control for perceptual effects, we contrasted the first (pure) inspection against the second, which can be made relative to first. Results revealed more activation in the striatum, amygdala, and ventromedial prefrontal cortex (vmPFC) during the naive inspections and in the anterior insula and dorsolateral prefrontal cortex (DLPFC) during the comparative inspections, showing that our coin inspection task activates known reward circuitry (Fig. 5A; Extended Data Table 5-1). The initial analysis also confirms that brain areas often associated with the default mode network (e.g., vmPFC) are more active during the naive inspection, while regions often associated with the task-positive response-selection network (e.g., DLPFC and insula) activate during the comparative inspection (Fox et al., 2005).

We then examined the specific effects of sequence valence and preference during each inspection separately and together (Extended Data Tables 5-2, 6-1). We contrasted preferred over rejected options regardless of option order, and this contrast revealed broad activation in the anterior insula, DLPFC, ACC, and VStr (Fig. 6A; Extended Data Table 5-2). Moreover, the more suboptimal decision-makers showed a deactivation in the right dorsal caudate nucleus (dCN) for inspection of growth versus decline (Fig. 5B; Extended Data Table 5-2) regardless of option order. In the comparative inspection there was an interaction between preference and option valence. During the comparative inspection subjects can begin to form a preference for or against the inspected option, and this second-order effect interacted with  $\tau$ . This effect had the direction that the suboptimal decision-makers showed bilateral hyperactivation in OFC/anterior insula and ventral pallidum when they preferred a growing rather than declining sequence (Fig. 6B; Extended Data Table 6-1).

The transition from experience to summary valuation was defined at the end of the inspection, which importantly was uncorrelated with the inspection onset because the coins streams varied widely in duration. Contrasting the end of the second inspection over the first, we found that the comparative summary was encoded bilaterally in putamen, caudate, vmPFC, lateral OFC, and cingulate cortex (Fig. 7A; Extended Data Table 7-1). This result seconds studies examining automatic computation of value in the absence of choice (Lebreton et al., 2009) and studies showing that subjective valuation engages medial PFC and striatum (Levy et al., 2011) with experienced value encoded in the anterior vmPFC (Smith et al., 2010). To investigate the role of the two distinct reward representations, we regressed brain activity at the end of both inspection epochs on GV and markdown. This analysis showed a considerable effect in the medial occipito-temporal area indicating the involvement of perceptual function in the inspection and analysis of the conditioned summary value. Moreover, bilateral activity in amygdala at the end of inspection was strongly correlated with GV in all participants, whereas in the suboptimal decision-makers, the markdown was strongly correlated with anterior insula activation. The result that only the suboptimal decision-makers encode the markdown is unsurprising given that there is very little variance in DM for high values of  $\tau$ . Furthermore, the more optimal the decision-maker, the stronger was GV encoded and the weaker was the markdown encoded (Fig. 7B,C; Extended Data Table 7-2).

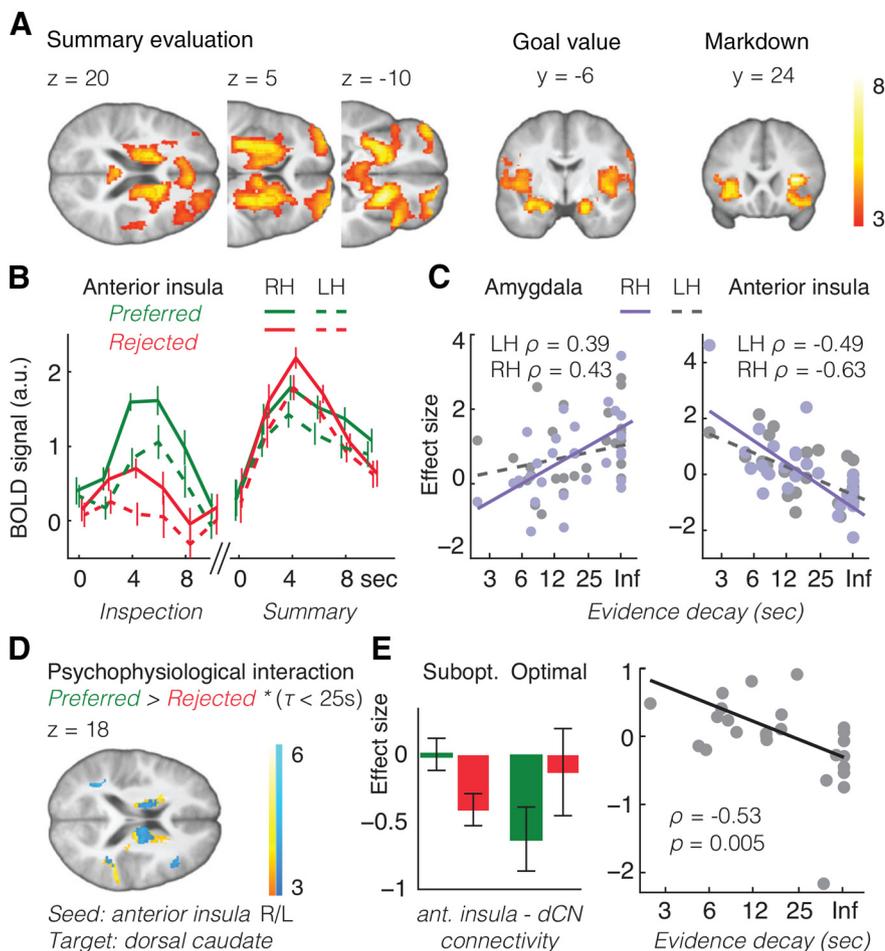
Because the task is self-paced, there may be a difference between participants in the rate at which they experienced the rewards. To address this issue, we calculated the reward rate in two ways: (1) GV experienced per minute (GV/min) defined as the sum of the individual GVs for the chosen options relative to the self-paced duration of the task, and (2) rewards experienced per minute (Rwd/min) defined as the number of rewards relative to the duration of the task. Results showed that GV/min was negatively correlated with  $\kappa$  ( $\rho^2 = 0.86$ ,  $p = 4.6 \times 10^{-12}$ ); that is, the more compressive was the value function the lower was GV/min (because GV is lower in these participants). GV/min and Rwd/min were also correlated ( $\rho^2 = 0.32$ ,  $p = 2.3 \times 10^{-3}$ ), but the correlation between Rwd/min and  $\kappa$  was insignificant. To analyze the effect of reward rate regardless of the effect of value function, we therefore included Rwd/min as covariate at the second level and repeated all the neuroimaging analyses. None of these analyses changed the main effects reported in Figures 5-7 or revealed any interaction with reward rate.

To identify areas involved in the modulation of the distinct reward representations, GV and DM, we computed the PPI of neural activity during the inspection in the amygdala and



**Figure 6.** Striatopallidal activation during valuation and time-resolved effects gradually building up over time after stimulus onset. **A**, Bilateral VStr encoding preference during inspection (Extended Data Table 5-2). **B**, Preference for growth encoded differentially in optimal and suboptimal decision-makers in bilateral anteroventral pallidum and VStr (Extended Data Table 6-1). The more suboptimal decision-makers exhibit a downregulation of the ventral pallidum when preferring a declining option, whereas the more optimal decision-makers show no different activity. The differential pallidum activity correlates with the decay ( $\tau$ ) in individual incentive-value models (excluding  $\tau = \infty$ ; when the two outliers are removed the correlation remains  $-0.47$ ,  $p = 0.018$ ).

anterior insula. Since we wanted to look for areas in the brain that down-modulate value responses for and lead to rejection of decreasing sequences, the PPI analyses look for changes in functional coupling as a function of sequence valence and preference. The PPI analysis for the anterior insula revealed strong functional coupling with a considerable portion of the dorsal caudate (Fig. 7D; Extended Data Table 7-3). The more suboptimal the decision-maker, the more negative was the correlation for the rejected option, whereas in the more optimal decision-makers this relationship held for the preferred option (Fig. 7E). The RH caudate cluster identified in the PPI for the right insula includes the caudate activation reported in Figure 5B. Moreover, for increasing sequences there was functional coupling between the anterior insula and some posterior-temporal and occipito-temporal areas in suboptimal decision-makers (Extended Data Table 7-3). The PPI analysis for the amygdala revealed functional coupling with the VStr. Brain activity in the anterior aspect of the pallidum, the antero-ventral caudate and putamen correlated with amygdala activation for rejected sequences and this effect was more so when the sequence was increasing (Extended Data Table 7-4). There was no statistically significant coupling for



**Figure 7.** Summary valuation. **A**, Comparative summary engages bilateral putamen, caudate, vmPFC (BA10/11), posterior cingulate and lateral premotor (BA6/8; Extended Data Table 7-1). Encoding of GV involves bilateral amygdala, and encoding of DM involves anterior insulae (Extended Data Table 7-2). **B**, Dynamics of anterior insula activation during inspection and summary valuation. **C**, The more optimal is the decision-maker, the stronger is GV encoded (LH:  $p = 0.044$ ; RH:  $p = 0.027$ ); encoding of the markdown showed the opposite pattern (LH:  $p = 0.011$ ; RH:  $p = 0.0005$ ). **D**, PPI (Extended Data Tables 7-3, 7-4). Differential preference-dependent coupling between anterior insula (left, cold; right, hot) and dorsal caudate in optimal and suboptimal decision-makers. **E**, Effect sizes for the RH PPI analysis in panel D shown for illustration purposes. The differential coupling between anterior insula and dorsal caudate correlates with the decay ( $\tau$ ) in individual incentive-value models (excluding  $\tau = \infty$ ; when the outlier is removed the correlation remains  $-0.51$ ,  $p = 0.008$ ).

sequence valence alone and no functional coupling between the anterior insula and the amygdala within the parameters of the factorial design.

These results show that incentives such as GV and growth, and disincentives such as the markdown, as well as preference are all encoded during the inspection. To investigate the extent to which these aspects of an experience also mediate the expression of preference, we analyzed brain activity at the time of choice and we examined the effect of  $\tau$  (Extended Data Table 8-1). We first analyzed the contrasts overt choice (free versus imperative preferred) and preferred versus rejected. The overt choice contrast indicates the actual choice itself separated from its execution, whereas the preferred versus rejected contrast indicates associative aspects of the options separated from the choice itself. Using this delineation, we found that the overt choice was encoded in cingulate cortex, DLPFC, and anterior insula (Fig. 8Ai). We then regressed the BOLD contrast between the preferred and rejected options on the value difference signals and found that GV was encoded in vmPFC (Fig. 8Aii). Finally, we examined interaction between brain activity at the time of choice (vs implicit baseline) and  $\tau$  and found that for the free choice the

more suboptimal the decision-maker, the stronger was activity in the DLPFC (Fig. 8B). In other words, although the difference in GV between the two options was encoded robustly in vmPFC, some decision-makers still make suboptimal choices, and the brain activity in the suboptimal decision-makers is characterized by increased recruitment of DLPFC as they reveal their preference in the free choice.

## Discussion

Our results reveal a diverse role for the insular cortex in summing up experience. The anterior insula contains so-called von Economo neurons abundant with dopamine receptors (Allman et al., 2005), and the distinct agranular frontoinsula is thought to be functionally related to the OFC (Morel et al., 2013). It has been proposed that the anterior insula integrates interoceptive and motivational information and that the right anterior insula is more involved in this process than the left insula (Craig, 2009). Furthermore, physiology studies have indicated that the human insula is part of the visceral nervous system encoding homeostatic state (Craig, 2003). Thus, the encoding in the anterior insula of a DM for a declining sequence of rewards may serve to inform intuitive decision-makers of their “gut feeling.”

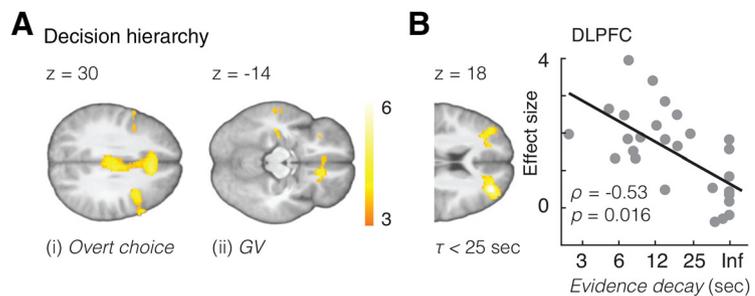
The insula is one of the most commonly activated regions in fMRI research and it can therefore be difficult to identify the specificity of its role. A pragmatic approach would be to note that the anterior insula is known to consistently and selectively engage with working memory and emotional tasks (Phan et al., 2002; Yarkoni et al., 2011), and to speculate that the dynamic anterior insula activation in our study therefore reflects selective recruitment of emotive and working memory functions as necessary to encode the IV of an experience. However, this perspective disregards other established roles of the anterior insula in encoding risk and uncertainty (Huettel et al., 2005; Kuhn and Knutson, 2005; Symmonds et al., 2010), attention switching (Phan et al., 2002), disgust (Calder et al., 2007), etc. Below we therefore focus on the robust encoding of the DM and the link between dorsal caudate and anterior insula.

Our results show that the neural underpinning of duration neglect is a separate encoding of two distinct reward representations, GV and DM (Fig. 7A,C). GV and IV share variance related to sequence duration, and we did not want to remove the effect of duration from GV. To isolate the common variance in GV, we considered the DM, from which the shared variance was removed in order that the markdown captured the extent to which a sequence declined (its “downness”) regardless of its duration. Using GV and DM as regressors we found that GV was encoded in the amygdala; it is the objective value of a sequence

including the effect of duration. By contrast, the activation in the anterior insula that correlates with the markdown may be because of the displeasing effect of decline regardless of outcome duration.

The mechanistic interpretation of the DM is potentially twofold: the  $\tau$  parameter can be seen to quantify either a leak in a memory integration process or a penalty imposed on the reward sequence based on its temporal configuration. We favor the latter because of its theoretical alignment with the computational model, it squares with the conceptual conflict between long and decreasing reward sequences and it is supported by the neuroimaging data. The theoretical foundation for the computational model is that people continually discount an experienced reward in relation to historical cumulative reward. In other words, the mechanism we examine is Equation 2 (not the solution to Eq. 2). The reason why humans and other animals discount perceived reward in relation to past events supports a theoretical problem related to the bounded dynamic range of the nervous system. When evaluating an experience of unknown future duration, the brain would need to constantly scale the cumulative reward in an adaptive fashion to encode the sum. Equation 2 proposes a simpler mechanism whereby summary evaluation for sequences of unknown continuing duration is possible without running the neural code out-of-bounds. Under uncertainty, people are known to rely on simple judgmental operations according to the so-called availability heuristic (Tversky and Kahneman, 1974). Our neural results are also aligned with the interpretation that the temporal configuration of a sequence imparts a penalty on the sequence; that penalty is the DM and it is encoded robustly in the anterior insula. If the markdown was merely a leak it is not obvious why the brain would retain a robust encoding of it. The fact that the markdown is encoded separately in the brain also supports the notion that it is not merely a distortion of the internal representation of GV. While the markdown in these experiments is a disincentive similar to a monetary acquisition cost, it is not incurred as a discount on the GV of the outcome; rather, it is a psychophysical construct which may well be experienced by decision-makers and which is encoded in the anterior insula of the brain

Wittmann et al. (2016) showed that the human brain can encode simultaneous representations related to historical rewards and foraging behavior. The DM in our study is also an index of the outcome's trend, and in light of that view, our results support the notion that outcome trends not only inform decision-making but also that they are encoded in neural circuitry separate from primary reward structures. We have thus demonstrated a neural dissociation between experienced values and reward trends. We have shown that the amygdala encodes the total GV of extended outcomes, whereas the anterior insulae encodes a DM penalizing temporary decline and leading to overvaluation of experience involving temporary growth (Fig. 7A,C). The engagement of the dorsal striatum during inspection suggests a tangible role for learning from experience (Dolan and Dayan, 2013), and the downregulation of dorsal caudate activation indicates that this process may be suspended in suboptimal decision-makers who are too favorably impressed by the experience of growth (Fig. 5B). This speculation could be taken to suggest that when the people who are most susceptible to a



**Figure 8.** Successive aspects of choice (Extended Data Table 8-1). **A**, Overt choice involves cingulate, DLPFC, and anterior insula (*i*; DLPFC shown in panel **B**, left, anterior insula not shown); GV encoded in vmPFC (*ii*). **B**, Differential recruitment of DLPFC during free choice in optimal and suboptimal decision-makers.

favorable appreciation of increasing outcomes observe a favorably impressive option, they suspend the commission of a deliberate analysis to memory. This effect seems to be more categorical and therefore not a direct neural index of the cumulative markdown (Fig. 2C). The differential coupling between the anterior insula and dorsal caudate (Fig. 7D) accords with these perspectives and suggests that suboptimal decision-makers may primarily encode aspects of the unwanted option whereas the more optimal decision-makers tend to encode aspects of the preferred option stronger (Fig. 7E).

Hare et al. (2009) showed that brain activity in vmPFC correlated with GVs, whereas activity in DLPFC correlated with the self-control that dieters exercise when they integrate competing concerns relating to healthiness and tastiness of foods. We found that brain activity in vmPFC correlated with the difference in GV between the preferred and rejected option (Fig. 8A), whereas activity in DLPFC was negatively correlated with an index of optimality (Fig. 8B). While it is well known that vmPFC encodes the GV difference signal (Boorman et al., 2009; Rushworth et al., 2011), DLPFC is part of a network that has been implicated in flexible control of attention to competing attributes (Leong et al., 2017; Tusche and Hutcherson, 2018). Thus, suboptimal decision-makers may be more likely to be switching attention between GV and DM and integrating both considerations, in a way that optimal decision-makers are not.

Previous studies have argued that amygdala activation mediates emotional responses that can lead to suboptimal behavior (De Martino et al., 2006; Roiser et al., 2009). While the amygdala is known to encode a wide range of signals relating to emotional experience, decision-making, and reward processing (Adolphs, 2010; Rutishauser et al., 2015), our results do not support the notion that suboptimal decision-making is underwritten by amygdala activity. The amygdala has also been shown to act as a reliable integrator of future rewards encoding the outcome of extended saving actions (Zangemeister et al., 2016), and our results are more compatible with a rational role for the amygdala in decision-making.

Although it is clearly suboptimal to discount declining sequences in the current experimental setting, in the real world and in other contexts, it might be optimal for the brain to attend to whether reward values are declining or increasing. The idea that increasing reward values signal that something better is coming up is ecologically plausible. Thus, reward contrasts may be honest indicators of the prospect for slowly varying events (Ossmy et al., 2013) serving as reliable signals optimizing fitness. According to this notion, there would be survival value in the repulsion to declining reward sequences. Such a mechanism concurs with the strong tendency of animals to approach stimuli

associated with rewards and to withdraw from stimuli associated with danger (Dickinson and Mackintosh, 1978). Therefore contrast-guided evaluation may be an ecologically beneficial strategy for future events. However, for retrospective valuations, an inclination in favor of persistent growth is clearly disadvantageous.

Taken together, our results show that suboptimal choice can be the result of robust neural representation of a displeasing aspect of the experience such as temporary decline. The study also challenges a popular belief that suboptimal decision-making is somehow routed in primitive neural structures whereas more astute reasoning is thought to emerge from the more evolved frontal executive system. Rather, we have here demonstrated that the summary value of an extended experience, which may be difficult and/or effortful to calculate, is encoded robustly in the amygdala, a brain region highly conserved across vertebrate evolution. By contrast, the markdown, a more intuitive construct, is encoded in the anterior insula, a paralimbic structure functionally related to the OFC. This neural correlate of the markdown may serve as a key discriminator between functional aspects of Kahneman's dichotomy of experienced utility and decision utility (Kahneman, 2003). We have shown that the human brain can encode these aspects in separate neural structures, and the participants in our study seem to have recruited this network differentially depending on how optimally they behaved.

## References

- Adolphs R (2010) What does the amygdala contribute to social cognition? *Ann NY Acad Sci* 1191:42–61.
- Allman JM, Watson KK, Tetreault NA, Hakeem AY (2005) Intuition and autism: a possible role for Von Economo neurons. *Trends Cogn Sci* 9:367–373.
- Ariely D (1998) Combining experiences over time: the effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *J Behav Decis Making* 11:19–45.
- Ashburner J (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* 38:95–113.
- Ashburner J, Friston KJ (2005) Unified segmentation. *Neuroimage* 26:839–851.
- Barbera S, Bossert W, Pattanaik PK (2004) Ranking sets of objects. In: *Handbook of utility theory: volume 2 extensions* (Barbera S, Hammond P, Seidl C, eds), pp 893–977. New York: Springer.
- Baumgartner H, Sujan M, Padgett D (1997) Patterns of affective reactions to advertisements: the integration of moment-to-moment responses into overall judgments. *J Marketing Res* 34:219–232.
- Becker GM, DeGroot MH, Marschak J (1964) Measuring utility by a single-response sequential method. *Behav Sci* 9:226–232.
- Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62:733–743.
- Calder AJ, Beaver JD, Davis MH, van Ditzhuijzen J, Keane J, Lawrence AD (2007) Disgust sensitivity predicts the insula and pallidal response to pictures of disgusting foods. *Eur J Neurosci* 25:3422–3428.
- Craig AD (2003) Interoception: the sense of the physiological condition of the body. *Curr Opin Neurobiol* 13:500–505.
- Craig AD (2009) How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci* 10:59–70.
- Cusack R, Vicente-Grabovetsky A, Mitchell DJ, Wild CJ, Auer T, Linke AC, Peelle JE (2014) Automatic analysis (aa): efficient neuroimaging workflows and parallel processing using Matlab and XML. *Front Neuroinform* 8:90.
- Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19:430–441.
- De Martino B, Kumaran D, Seymour B, Dolan RJ (2006) Frames, biases, and rational decision-making in the human brain. *Science* 313:684–687.
- Dickinson A, Mackintosh NJ (1978) Classical conditioning in animals. *Annu Rev Psychol* 29:587–612.
- Do AM, Rupert AV, Wolford G (2008) Evaluations of pleasurable experiences: the peak-and-end rule. *Psychon B Rev* 15:96–98.
- Dolan RJ, Dayan P (2013) Goals and habits in the brain. *Neuron* 80:312–325.
- Donoso M, Collins AG, Koechlin E (2014) Human cognition. *Foundations of human reasoning in the prefrontal cortex. Science* 344:1481–1486.
- Eklund A, Nichols TE, Knutsson H (2016) Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA* 113:7900–7905.
- Elliott R, Newman JL, Longe OA, Deakin JF (2003) Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23:303–307.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
- Frederick S, Loewenstein G (2008) Conflicting motives in evaluations of sequences. *J Risk Uncertain* 37:221–235.
- Fredrickson BL (2000) Extracting meaning from past affective experiences: the importance of peaks, ends, and specific emotions. *Cogn Emot* 14:577–606.
- Fredrickson BL, Kahneman D (1993) Duration neglect in retrospective evaluations of affective episodes. *J Pers Soc Psychol* 65:45–55.
- Haber SN (2011) Neuroanatomy of reward: a view from the ventral striatum. In: *Neurobiology of sensation and reward* (Gottfried JA, ed). Boca Raton: CRC Press.
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28:5623–5630.
- Hare TA, Camerer CF, Rangel A (2009) Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* 324:646–648.
- Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* 117:500–544.
- Huettel SA, Song AW, McCarthy G (2005) Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J Neurosci* 25:3304–3311.
- Kahneman D (2003) Maps of bounded rationality: psychology for behavioral economics. *Am Econ Rev* 93:1449–1475.
- Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. *J Neurosci* 25:4806–4812.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540.
- Kuhnen CM, Knutson B (2005) The neural basis of financial risk taking. *Neuron* 47:763–770.
- Kuo WJ, Sjoström T, Chen YP, Wang YH, Huang CY (2009) Intuition and deliberation: two systems for strategizing in the brain. *Science* 324:519–522.
- Lazzaro SC, Rutledge RB, Burghart DR, Glimcher PW (2016) The impact of menstrual cycle phase on economic choice and rationality. *PLoS One* 11:e0144080.
- 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.
- Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y (2017) Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* 93:451–463.
- Levy I, Lazzaro SC, Rutledge RB, Glimcher PW (2011) Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J Neurosci* 31:118–125.
- Liu X, Hairston J, Schrier M, Fan J (2011) Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 35:1219–1236.
- Morel A, Gallay MN, Baechler A, Wyss M, Gallay DS (2013) The human insula: architectonic organization and postmortem MRI registration. *Neuroscience* 236:117–135.
- Morrison SE, Salzman CD (2010) Re-valuing the amygdala. *Curr Opin Neurobiol* 20:221–230.
- Mumford JA, Poline JB, Poldrack RA (2015) Orthogonalization of regressors in fMRI models. *PLoS One* 10:e0126255.

- Nurmi H (2006) Models of political economy. Boca Raton: Taylor and Francis.
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4:95–102.
- Ossmy O, Moran R, Pfeffer T, Tsetsos K, Usher M, Donner TH (2013) The timescale of perceptual evidence integration can be adapted to the environment. *Curr Biol* 23:981–986.
- Paton JJ, Belova MA, Morrison SE, Salzman CD (2006) The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439:865–870.
- Phan KL, Wager T, Taylor SF, Liberzon I (2002) Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16:331–348.
- Redelmeier DA, Kahneman D (1996) Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain* 66:3–8.
- Rich EL, Wallis JD (2014) Medial-lateral organization of the orbitofrontal cortex. *J Cogn Neurosci* 26:1347–1362.
- Robbins TW, Everitt BJ (1996) Neurobehavioural mechanisms of reward and motivation. *Curr Opin Neurobiol* 6:228–236.
- Roiser JP, de Martino B, Tan GC, Kumaran D, Seymour B, Wood NW, Dolan RJ (2009) A genetically mediated bias in decision making driven by failure of amygdala control. *J Neurosci* 29:5985–5991.
- Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70:1054–1069.
- Rutishauser U, Mamelak AN, Adolphs R (2015) The primate amygdala in social perception - insights from electrophysiological recordings and stimulation. *Trends Neurosci* 38:295–306.
- Schreiber CA, Kahneman D (2000) Determinants of the remembered utility of aversive sounds. *J Exp Psychol Gen* 129:27–42.
- Slotnick SD, Moo LR, Segal JB, Hart J Jr (2003) Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. *Brain Res Cogn Brain Res* 17:75–82.
- Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J Neurosci* 30:2490–2495.
- Stevens SS (1957) On the psychophysical law. *Psych Rev* 64:153–181.
- Symmonds M, Bossaerts P, Dolan RJ (2010) A behavioral and neural evaluation of prospective decision-making under risk. *J Neurosci* 30:14380–14389.
- Taylor JR, Williams N, Cusack R, Auer T, Shafto MA, Dixon M, Tyler LK, Cam C, Henson RN (2017) The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* 144:262–269.
- Tsetsos K, Chater N, Usher M (2012) Salience driven value integration explains decision biases and preference reversal. *Proc Natl Acad Sci USA* 109:9659–9664.
- Tsetsos K, Moran R, Moreland J, Chater N, Usher M, Summerfield C (2016) Economic irrationality is optimal during noisy decision making. *Proc Natl Acad Sci USA* 113:3102–3107.
- Tusche A, Hutcherson CA (2018) Cognitive regulation alters social and dietary choice by changing attribute representations in domain-general and domain-specific brain circuits. *Elife* 7.
- Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. *Science* 185:1124–1131.
- Vestergaard MD, Schultz W (2015) Choice mechanisms for past, temporally extended outcomes. *Proc R Soc B* 282:2014–1766.
- Vickrey W (1961) Counterspeculation, auctions, and competitive sealed tenders. *J Fin* 16:8–37.
- Von Neumann J, Morgenstern O (1947) Theory of games and economic behavior, Ed 2. Princeton: Princeton University Press.
- Wittmann MK, Kolling N, Akaishi R, Chau BK, Brown JW, Nelissen N, Rushworth MF (2016) Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat Commun* 7:12327.
- Woo CW, Krishnan A, Wager TD (2014) Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 91:412–419.
- Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8:665–670.
- Zangemeister L, Grabenhorst F, Schultz W (2016) Neural basis for economic saving strategies in human amygdala-prefrontal reward circuits. *Curr Biol* 26:3004–3013.