nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-019-0589-3

# Diverse motives for human curiosity

Kenji Kobayashi[1]*, Silvio Ravaioli[2,3], Adrien Baranès[4], Michael Woodford[3] and Jacqueline Gottlieb[1,4,5]

**Curiosity—our desire to know—is a fundamental drive in human behaviour, but its mechanisms are poorly understood. A classical question concerns the curiosity motives. What drives individuals to become curious about some but not other sources of information?[1] Here we show that curiosity about probabilistic events depends on multiple aspects of the distribution of these events. Participants ($n=257$) performed a task in which they could demand advance information about only one of two randomly selected monetary prizes that contributed to their income. Individuals differed markedly in the extent to which they requested information as a function of the ex ante uncertainty or ex ante value of an individual prize. This heterogeneity was not captured by theoretical models describing curiosity as a desire to learn about the total rewards of a situation[2,3]. Instead, it could be explained by an extended model that allowed for attribute-specific anticipatory utility—the savouring of individual components of the eventual reward—and postulates that this utility increased nonlinearly with the certainty of receiving the reward. Parameter values fitting individual choices were consistent for information about gains or losses, suggesting that attribute-specific anticipatory utility captures fundamental heterogeneity in the determinants of curiosity.**

Following a wave of research that peaked in the late twentieth century and subsequently waned[4], a resurgence of interest in the mechanisms of curiosity has been motivated by increased appreciation of its importance for brain and cognitive function[5–8]. Many recent studies of curiosity rely on so-called non-instrumental paradigms, in which participants can demand advance information about a future outcome (punishment or reward) but cannot take actions to exploit the information they sample. These studies have shown that humans and other animals have reliable preferences for non-instrumental information[9–11], and these preferences are encoded in neural systems of reward and motivation[12,13], and impact memory, attention and gaze[5,14–16].

The robust demand for non-instrumental information revealed by these studies poses significant challenges to traditional decision theories, in which information value is defined in terms of reward gains. The theories allow for the fact that decision-makers make tradeoffs between exploration and exploitation by seeking to reduce uncertainty on immediate time scales to maximize the rewards they obtain on longer time scales[17], but they cannot explain the desire to obtain information as a good in itself, independent of instrumental incentives. The prevalence of curiosity in behaviour therefore implies that decision-makers assign intrinsic value to some property (or properties) of internal states that are engendered by information.

Two prominent lines of work in decision theory provide potential answers about what these properties may be. One theoretical approach proposes that individuals prefer to hold more or less accurate beliefs about future outcomes (have preferences over the timing of resolution of uncertainty) independent of their preferences over the outcomes themselves[2]. A different approach proposes instead that information choice reflects a utility that individuals derive from anticipation—the desire to feel good by anticipating (savouring) positive outcomes, but avoid the dread associated with anticipating negative outcomes. Anticipatory utility is also proposed to be distinct from the utility of the outcomes themselves, and has been formalized in models of economic utility[18,19] and, more recently, in the reinforcement learning literature[3].

To date, these two theories have not been empirically contrasted and it is not known which one better describes human curiosity. An outstanding question pertains to cases in which decision-makers must select between competing sources of information that are relevant to a situation. Such cases are the rule in natural behaviour, in which decision-makers contend with complex situations that have multiple relevant features (or attributes) and it is not feasible to become fully informed about all of the features. Although the question of information selection has been long recognized as being key for curiosity[1], this question has been typically eschewed in laboratory settings, which have instead used tasks in which decision-makers receive rewards from a single source and have the opportunity to obtain perfect information about that sole source. It remains unknown to what extent current computational models and empirical studies capture how decision-makers choose which information to sample when they selectively interrogate a multi-attribute situation.

To examine this question, we tested participants on a task in which they received probabilistic payoffs from two independent sources (two randomly drawn monetary prizes) but could request information about only one source. On each trial, participants were shown two distributions (lotteries), defined by a mean and variance, and were told that the computer will draw one prize from each distribution (randomly with uniform probability) and pay out the sum of the prizes (Fig. 1a). Participants were not by default informed about the precise value of the prizes that happened to be drawn, but were instead asked to choose one prize whose value they wished to reveal. In the example in Fig. 1a, the participant inquired about the left lottery and learned that the prize from that lottery was 240 points, but remained ignorant of the precise value of the additional prize from the right lottery. At the end of the block, participants received a monetary payoff equal to the sum of the prizes they drew on one trial that was randomly selected from those they had played. Thus, the participants' rewards were determined by chance, and the information they gathered was non-instrumental.

[1]The Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA. [2]Sant'Anna School of Advanced Studies, Pisa, Italy. [3]Department of Economics, Columbia University, New York, NY, USA. [4]Department of Neuroscience, Columbia University, New York, NY, USA. [5]The Kavli Institute for Brain Science, Columbia University, New York, NY, USA. *e-mail: kenji.kobayashi@berkeley.edu

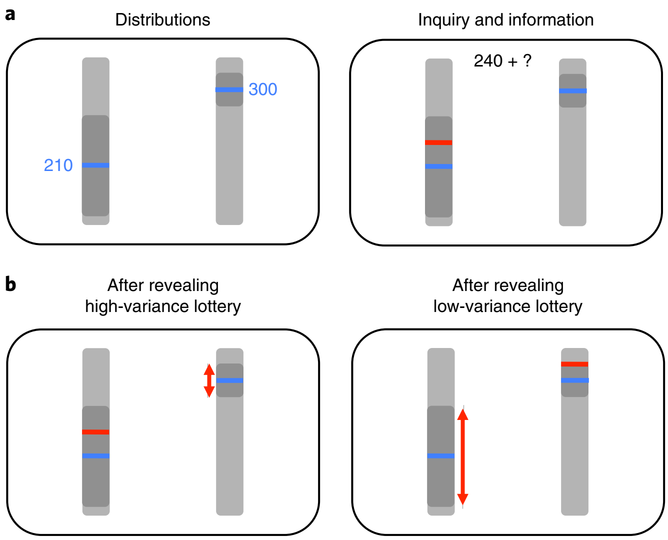


**Fig. 1 | The task. a**, The choice screen on each trial depicted two lotteries that differed in their variance (represented by the length of dark grey bar) and EV (midpoint of the bar, marked by the blue line and numerical value). The lighter grey background rectangle indicated the total range of points possible in the experiment (0–500 points, constant in all trials, shown here not to scale). After indicating their choice of which lottery to observe, the participants received immediate feedback about the precise prize that had been drawn from the chosen lottery in the form of a horizontal red bar and a numerical value displayed at the top of the screen (right, illustrating the case in which the participant inquired about the left lottery). Larger numbers (at higher positions on the screen) indicated larger values of gains or losses. In this example, the lottery on the right had the higher potential gain on average in the gain domain version but the higher potential loss in the loss domain version. **b**, The reduction of uncertainty about the total outcome depends on the variance, but not EV, of the inspected lottery. After revealing the value of a prize from the high-variance lottery (left), the decision-maker's remaining uncertainty (red arrows) is lower than when the prize of low-variance lottery is revealed (right). The remaining uncertainty does not depend on the revealed prize (height of the horizontal red line); hence, nor does it depend on the EV of the inspected lottery.

We independently manipulated the expected value (EV) and variance (uncertainty) of the distributions generating the prizes to determine which factor more strongly influenced the participants' choices (Methods). Importantly, this manipulation allowed us to differentiate between two potential motives for curiosity—the desire to reduce uncertainty about the total reward of the trial versus the desire to obtain information about an individual prize. As shown graphically in Fig. 1b, participants could expect that their residual uncertainty about the total reward after the information would be equal to the uncertainty of the unrevealed prize (Fig. 1b). Thus, their uncertainty about the sum of the prizes would be minimized if they inquired about the lottery with the larger ex ante uncertainty, regardless of the EV of this lottery. We show formally that both the theory of Kreps and Porteus[2] and the reinforcement learning model of Iigaya et al.[3] assume that the value of non-instrumental information depends strictly on the extent to which the information resolves uncertainty about the total future utility, and thus predict that curiosity in this task would be strictly a function of the uncertainty of a lottery (Supplementary Notes 1 and 3).

In contrast with this prediction, participants showed both a sensitivity to lottery uncertainty and a prominent bias to inspect the lottery with the higher EV, as we describe in detail below. To capture these observations, we devised a computational model that allows information demand to be motivated not only by uncertainty about the total utility but also by a form of anticipatory utility, which we call attribute-specific anticipatory utility (Supplementary Notes 1 and 2). This model allows for the possibility that individuals derive utility from advance information about individual components of a total reward (that is, a direct effect of the information that is received). Specifically, our model postulates that the degree of savouring of rewards that one is already certain to receive is greater than the savouring of rewards that are only possible, but not certain, by a factor larger than the increase in probability (Supplementary Note 2).

Our model derivations showed that the weights of each motive can be recovered from the parameters of a simple psychometric choice function that plots the probability of inspecting the high-variance lottery as a function of its EV relative to the EV of the low-variance lottery ($\Delta$EV) (equation (1) and Supplementary Note 1, equation (8)). We thus fitted choice functions using logistic regression and recovered the parameter $w_{\mathrm{var}}$, which characterizes a participant's desire to reduce uncertainty about the total outcome (the vertical shift of the choice function), and parameter $w_{\Delta\mathrm{EV}}$, which characterizes a participant's sensitivity to attribute-specific anticipatory utility (the slope of the choice function).

Across the population, parameters $w_{\mathrm{var}}$ and $w_{\Delta\mathrm{EV}}$ were both significantly larger than 0, indicating that people were strongly sensitive to both motives (Fig. 2). When sampling information about gains, parameter $w_{\mathrm{var}}$ was positive, indicating an overall preference for the early resolution of uncertainty about the total outcome (Fig. 2a, ordinate; median = 1.53; mean = 1.99; s.e.m. = 0.13; Wilcoxon signed rank test, $Z=11.89$; $P<0.001$; $r$ (the $Z$ value divided by the square root of the number of participants) = 0.74; 95% confidence interval (CI) = 0.68–0.79; $n=257$). Parameter $w_{\Delta\mathrm{EV}}$ was also larger than 0, indicating that people were significantly motivated to view the higher-value individual prize, as predicted by attribute-specific anticipatory utility (Fig. 2a, abscissa; median = 3.32; mean = 3.57; s.e.m. = 0.26; $Z=11.55$; $P<0.001$; $r=0.72$; 95% CI = 0.64–0.78).

Similarly, when participants sampled information about losses, they showed a preference for early resolution of uncertainty (Fig. 2b, ordinate; $w_{\mathrm{var}}$ median = 0.90; mean = 1.55; s.e.m. = 0.19; $Z=6.77$; $P<0.001$; $r=0.57$; 95% CI = 0.42–0.66; $n=140$) and for viewing the higher-value individual prize (Fig. 2b, abscissa; $w_{\Delta\mathrm{EV}}$ median = 0.96; mean = 1.59; s.e.m. = 0.37; $Z=4.30$; $P<0.001$; $r=0.36$; 95% CI = 0.20–0.50). The fact that the parameter $w_{\Delta\mathrm{EV}}$ was positive shows that participants were primarily interested in the lottery that had the better outcome (smaller individual loss); that is, they were motivated to avoid dread from anticipating a negative outcome rather than to obtain information about the most salient outcome (the largest individual loss[20]). Because the more desirable lottery was lower on the screen in the loss condition (but higher on the screen in the gain condition), these results also rule out a trivial strategy of simply inquiring about the lottery at a fixed position in the visual field.

In the subset of participants who completed tasks in both gain and loss domains, the parameters $w_{\mathrm{var}}$ and $w_{\Delta\mathrm{EV}}$ were highly correlated, indicating that participants followed similar sampling strategies for losses and gains (Fig. 2c,d; Spearman's rho for $w_{\mathrm{var}}=0.65$; $P<0.001$; rho for $w_{\Delta\mathrm{EV}}=0.49$; $P<0.001$; $n=140$). However, we noticed that $w_{\Delta\mathrm{EV}}$ in the loss condition was smaller than in the gain condition (Fig. 2d; $Z=4.48$; $P<0.001$; $r=0.38$; 95% CI = 0.22–0.52) and, in a subgroup of participants who performed the loss condition before the gain condition, $w_{\Delta\mathrm{EV}}$ in the loss condition failed to reach statistical significance (Supplementary Table 1; median = $1.71\times10^{-7}$; mean = 0.67; s.e.m. = 0.77; $Z=0.48$; $P=0.63$; $r=0.07$; 95% CI = −0.25–0.36; $n=43$). Thus, dread from anticipating a negative outcome may be more sensitive to contextual factors relative to savouring a positive outcome, but attribute-specific anticipatory utility related to savouring and dread share significant common variability.

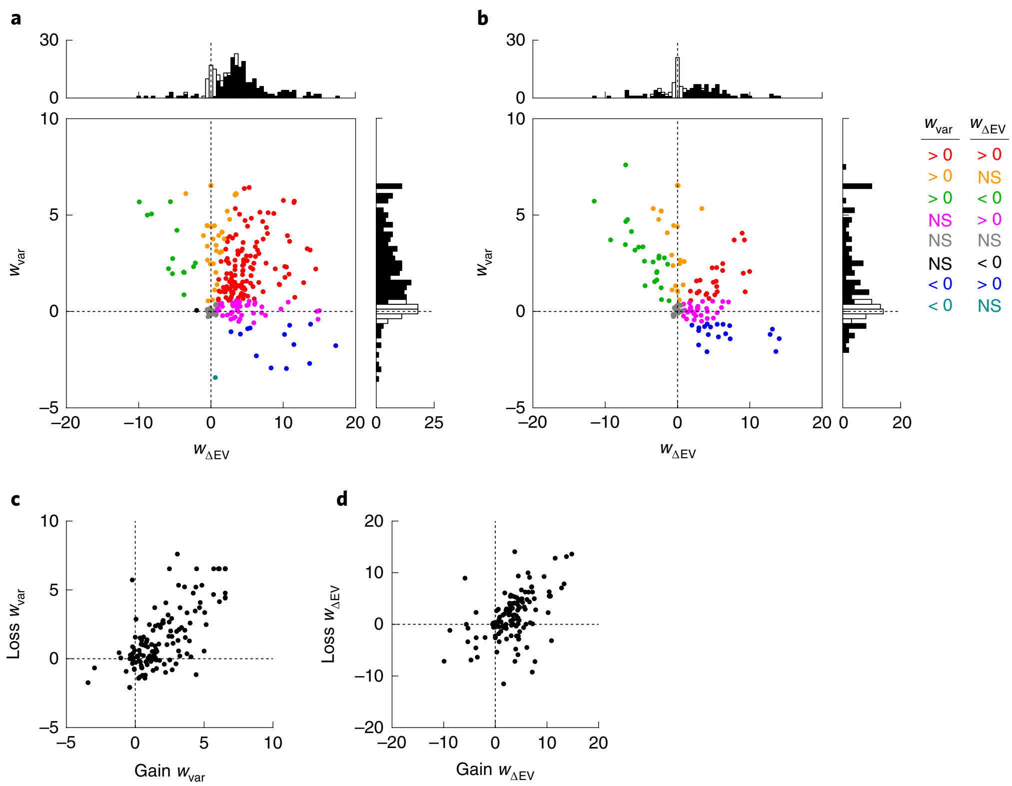


**Fig. 2 | Decision weights in the gain and loss domains. a,b**, Joint distributions of $w_{\Delta\mathrm{EV}}$ and $w_{\mathrm{var}}$ in the gain (**a**) and loss domains (**b**). Each point represents one participant, and the colour indicates the sign and significance ($\alpha = 0.05$) of the parameter estimates. NS, not significant. Black bars in the marginal distributions show individually significant coefficients. **c,d**, Relationship between the parameter estimates $w_{\mathrm{var}}$ (**c**) and $w_{\Delta\mathrm{EV}}$ (**d**) for participants who were tested in both the gain and loss domains. Each point represents one participant.

Despite the clear trends at the population level, behaviour showed considerable individual variability. To illustrate this variability, we plotted each participant's fitted choice function in the gain condition in Fig. 3a and in the loss condition in Fig. 3b (grouping individuals based on the significance and sign of $w_{\Delta\mathrm{EV}}$ and $w_{\mathrm{var}}$, evaluated at the individual level by non-parametric permutations, for ease of visualization without implying distinct behavioural categories).

A minority of participants showed information demand that was independent of ΔEV ($w_{\Delta\mathrm{EV}}$ did not significantly deviate from 0; $P > 0.05$), as would be predicted by the Kreps and Porteus[2] and Iigaya et al.[3] theories (Fig. 3, top row). Of this subgroup, the vast majority preferred the early resolution of uncertainty ($w_{\mathrm{var}} > 0$; $P < 0.05$; orange), only one participant preferred late resolution in the gain domain ($w_{\mathrm{var}} < 0$; $P < 0.05$; dark cyan) and a few others were indifferent between early and late resolution ($w_{\mathrm{var}}$ did not significantly deviate from 0; $P > 0.05$; grey).

However, the majority of participants showed significant sensitivity to attribute-specific anticipatory utility (Fig. 3, second to fourth rows). Most participants in this subgroup preferred viewing the high-EV lottery, consistent with a desire to savour the higher-value individual prize (positive slopes; $w_{\Delta\mathrm{EV}} > 0$; $P < 0.05$; red, pink and blue). The influence of ΔEV was often accompanied by variable degrees of interest in the high-uncertainty lottery, showing that individual participants could be sensitive to both motives (red: $w_{\mathrm{var}} > 0$; $P < 0.05$; pink: $w_{\mathrm{var}}$ did not significantly deviate from 0; $P > 0.05$; blue: $w_{\mathrm{var}} < 0$; $P < 0.05$). Interestingly, no participant was preferentially interested in lotteries that had both a lower variance and lower EV; there are no curves that decline as a function of ΔEV ($w_{\Delta\mathrm{EV}} < 0$) and intercept the $x$ axis at negative values ($w_{\mathrm{var}} < 0$) in Fig. 3 and, correspondingly, there are no points in the lower left quadrant in Fig. 2a,b.

In summary, whether sampling information about gains or losses, people are motivated by attribute-specific anticipatory utility—the desire to savour (or avoid the dread of) individual components of their outcomes—and this motive can coexist with a desire to reduce uncertainty about the total outcome.

We conducted several analyses to establish that these findings are robust and not explained by spurious factors related to the task design or instructions. First, the results were replicated in several groups of participants who performed the task in a different order and were tested in two laboratory settings, as well as online, showing that they were highly robust with respect to the experimental setting (Supplementary Table 1).

Second, while one might suppose that the relative strength of the two motives would depend on the absolute values of the lotteries, parameter estimates were highly consistent between trials in which the EV of the high-variance lottery was higher or lower than the median, suggesting that individual strategies were robust across absolute values in the range used in our paradigm (Supplementary Fig. 1).

Third, one might be concerned that because our participants did not receive operant incentives for their choices, they may have adopted spurious or arbitrary choice strategies. This possibility is refuted by the fact that most participants had at least one significant coefficient, meaning that they based their strategy consistently on ΔEV or uncertainty; very few participants were insensitive to both factors (14/257 for gains and 16/140 for losses; grey in Figs. 2 and 3), as would be predicted by a random strategy. As additional verification of this claim, we analysed choice reaction times, reasoning that

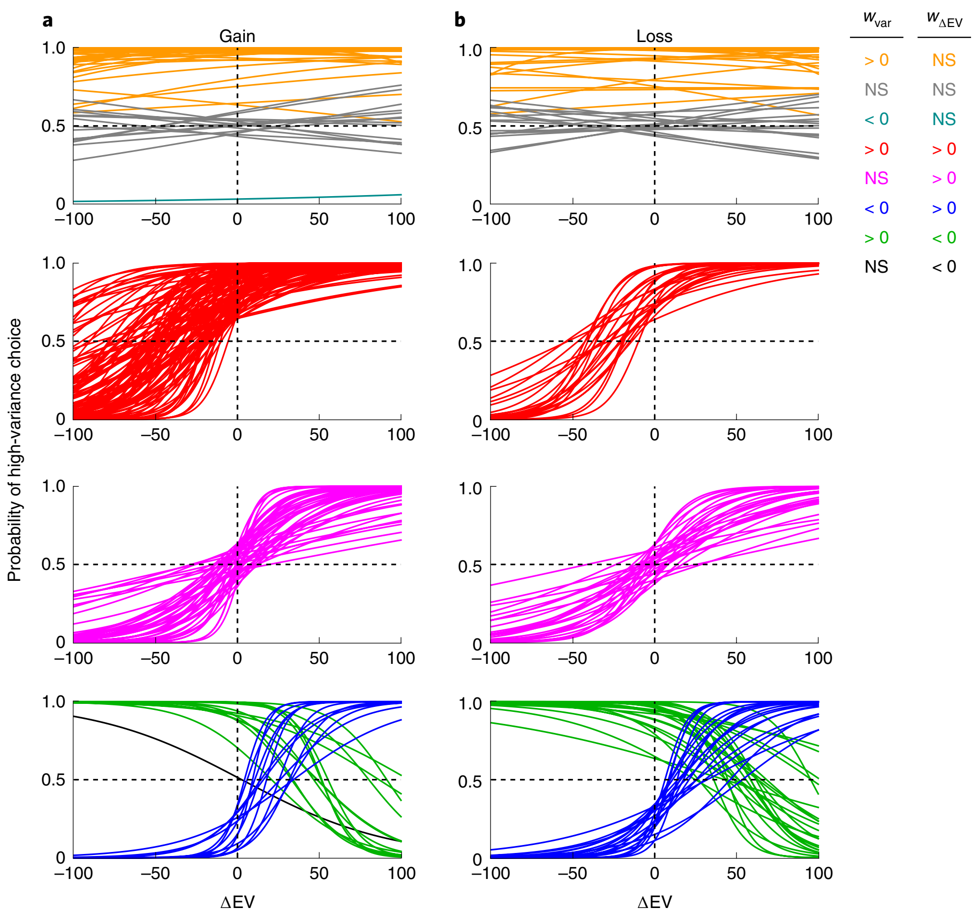


**Fig. 3 | Individual choice curves for observing decisions. a**, Gain domain. **b**, Loss domain. Each curve is the logistic function fit for one participant (equation (1)). Participants are grouped for ease of presentation based on the sign and statistical significance of $w_{var}$ and $w_{\Delta EV}$. The numbers of participants are as follows (total: 257 for gains; 140 for losses). Top, orange: 43 (17%) for gains; 29 (21%) for losses; grey: 14 (5%) for gains; 16 (11%) for losses; dark cyan: 1 (<1%) for gains; 0 for losses. Second row, red: 127 (49%) for gains; 24 (17%) for losses. Third row, pink: 45 (18%) for gains; 28 (20%) for losses. Bottom, blue: 13 (5%) for gains; 19 (14%) for losses; green: 13 (5%) for gains; 24 (17%) for losses; black: 1 (<1%) for gains; 0 for losses.

if participants carefully considered their choices, they should show the longest reaction times when the decision alternatives had the highest subjective similarity (that is, at the midpoint of the psychometric function for those who were sensitive to ΔEV; equation (2)). Reaction times were well fit by the model predictions, confirming that participants spent longer reaching decisions that had higher difficulty, consistent with a deliberative strategy (Fig. 4a,b; second row (participants with $w_{\Delta EV} > 0$ and $w_{var} > 0$): coefficient of choice difficulty on reaction time in gain = 1.68; 95% CI = 1.46–1.91; coefficient in loss = 2.25; 95% CI = 1.77–2.73; third row (participants with $w_{\Delta EV} > 0$ and $w_{var}$ = NS): coefficient in gain = 3.36; 95% CI = 2.98–3.75; coefficient in loss = 3.09; 95% CI = 2.41–3.78; fourth row (participants with $w_{\Delta EV} > 0$ and $w_{var} < 0$): coefficient in gain = 3.09; 95% CI = 2.60–3.58; coefficient in loss = 2.49; 95% CI = 2.05–2.93; fifth row (participants with $w_{\Delta EV} < 0$ and $w_{var} > 0$): coefficient in gain = 1.54; 95% CI = 1.00–2.07; coefficient in loss = 1.42; 95% CI = 1.00–1.83; all: $P < 0.001$).

Fourth, we considered another concern: that participants who were positively sensitive to ΔEV may have misunderstood the instructions and erroneously believed that their choices determined their payoffs in the two-lottery task. This possibility is unlikely given that participants explicitly confirmed their understanding in the instruction phase (Methods), and anticipatory utility was evident even in participants who performed the observing task only before other instrumental (incentivized) conditions (Supplementary Table 1, laboratory II).

As a further evaluation, we examined behaviour on the willingness-to-pay (WTP) task, in which the participants chose a single lottery that would contribute to their payoffs, and could trade off points in exchange for advance information about the prize drawn from that lottery (Methods). We reasoned that, if participants erroneously believed that their preference for the low-EV lottery would reduce their payoffs on the observing task (as was the case in the WTP task), they should show similar points of subjective equality (PSEs; the point at which the choice probability is 0.5) in both tasks. PSEs on the WTP task were significantly negative across the population, confirming that participants were willing to pay to obtain information (Supplementary Fig. 2; median = −26.8; $Z = 2.17$; $P = 0.030$; $r = 0.23$; 95% CI = 0.02–0.42; $n = 90$). However, this willingness to pay for information was quite small, although statistically significant, and the PSEs in the two-lottery task were much more negative relative to those in the WTP task (Supplementary Fig. 2; $Z = 5.53$; $P < 0.001$; $r = 0.58$; 95% CI = 0.40–0.70; median PSE in the observing tasks = −1.26). Therefore, participants understood the differences between tasks and were much more willing to express their interest in the uncertain prize in the observing task, in which there was no monetary cost to demanding information.

Lastly, we considered the possibility that participants adopted mixed strategies, choosing consistently with the task in some trials but falling back on an instrumental strategy on others. Quantitative model comparisons showed that such a mixed strategy model

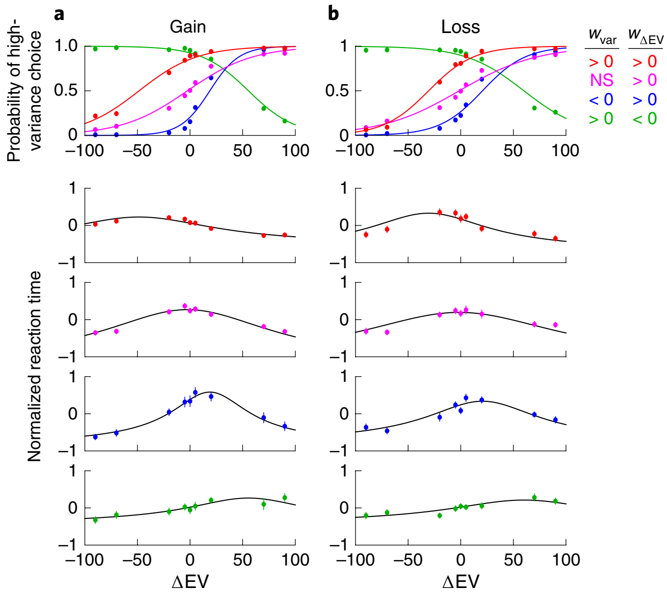


**Fig. 4 | Correspondence between reaction time and choices. a**, Gain domain. **b**, Loss domain. Top, average choices (lines: group-level fits) for the participants who were sensitive to ΔEV, grouped based on the sign and statistical significance of $w_{var}$ and $w_{\Delta EV}$. The subsequent rows show the population average of *z*-scored reaction times for each group (error bars represent s.e.m.). In all groups, reaction times were longer with more difficult choices (lines: model fits of equation (2)); the coefficient of the choice difficulty term was always positive (second row, gain: $F(1, 1{,}141) = 209.93$; coefficient = 1.68; 95% CI = 1.46–1.91; loss: $F(1, 214) = 86.13$; coefficient = 2.25; 95% CI = 1.77–2.73; third row, gain: $F(1, 403) = 295.35$; coefficient = 3.36; 95% CI = 2.98–3.75; loss: $F(1, 250) = 79.97$; coefficient = 3.09; 95% CI = 2.41–3.78; fourth row, gain: $F(1, 115) = 154.25$; coefficient = 3.09; 95% CI = 2.60–3.58; loss: $F(1, 169) = 125.11$; coefficient = 2.49; 95% CI = 2.05–2.93; fifth row, gain: $F(1, 115) = 31.98$; coefficient = 1.54; 95% CI = 1.00–2.07; loss: $F(1, 214) = 44.65$; coefficient = 1.42; 95% CI = 1.00–1.83; all: $P < 0.001$).

provided a better fit relative to the original model for only a handful of participants (Supplementary Note 4; 14/257 for the gain domain and 7/140 for the loss domain), and a reanalysis of the data excluding these participants produced equivalent results at the population level (Supplementary Table 1).

Having established the robustness of our findings, we performed two additional analyses to determine whether information seeking on this task was related to other behavioural metrics. First, we examined whether individual variations in sensitivity to uncertainty and anticipatory utility were correlated with corresponding effects of ΔEV and uncertainty in the conventional risk-taking paradigm (Methods). Participants showed much larger effects of uncertainty in the observing task relative to the risk task, and conversely, showed larger effects of ΔEV in risk taking relative to observing (Supplementary Fig. 3), supporting the conclusion that they distinguished between instrumental and non-instrumental contingencies. Moreover, parameter estimates were uncorrelated across the tasks, except for a mild correlation in the sensitivity to uncertainty in the gain domain (Supplementary Fig. 3), suggesting that people adopted largely distinct task-specific strategies.

Second, we administered to a subset of participants questionnaires assessing behavioural traits including avoidance/approach behaviours (behavioural inhibition system (BIS) and behavioural approach system (BAS) scores[21]), anticipatory and consummatory aspects of pleasure[22], obsessive–compulsive traits[23], anxiety[24], depression[25], curiosity/sensation seeking[26,27] and real-world domain-specific risk taking (DOSPERT)[28]. An automatic variable selection procedure using Lasso regularization (Methods) showed that none of these scales (or demographic data related to age, gender and education) was associated with the sensitivity to uncertainty or ΔEV in the gain domain, nor with the sensitivity to uncertainty in the loss domain. However, the sensitivity to ΔEV in the loss domain did show several relationships, including negative association with the BIS score, positive association with the BAS score, negative association with real-world risk-taking tendency, and tendencies to be larger in men relative to women and in undergraduate relative to graduate students (Supplementary Table 2). This suggests that dread may be motivated by an approach to positive items (rather than worry about negative outcomes) and may be associated with risk-avoidant attitudes. However, while these factors emerged in the variable selection procedure, they only showed mild correlations in individual analyses (Supplementary Fig. 4), suggesting that they explain limited variation of behaviour in our task.

We showed that the interest in non-instrumental information is shaped by two motives related, respectively, to holding accurate beliefs about the total future utility and learning about individual attributes of a situation. These motives can be computationally distinguished, but they jointly shape non-instrumental information demand, combining with different strengths in different individuals to produce heterogeneous strategies. Our findings thus go beyond recent accounts that tend to portray curiosity as a homogeneous process (for example, ref. [7]) and instead show that it entails considerable heterogeneity.

As in previous studies of non-instrumental information demand, our task is clearly distinguished from the exploration–exploitation literature by the absence of instrumental incentives. In an exploration–exploitation scenario, the decision-makers seek to maximize operant gains, and exploration is a priori considered as motivated by reward maximization on longer time scales. In our task, in contrast, participants had no control over the rewards they obtained and could not exploit the information they sampled on any time scale. Because of this feature, models of exploration and exploitation (such as traditional models of economic choice) predict that behaviour in our task would be random with no consistent demand for non-instrumental information, and cannot account for our findings that the vast majority of participants have non-random, well-defined informational strategies.

Our conclusion that many participants made observing decisions as though they were seeking to reduce uncertainty about their total reward is consistent with recent empirical findings by van Lieshout et al.[29], as well as with theoretical frameworks such as the free-energy principle[30] and models of non-instrumental information demand[2,3]. However, our data show that participants had an additional, distinct drive to reduce their uncertainty about specific components of their total reward, requiring an extension of these theoretical models.

An important comparison is between our study and the model proposed by Iigaya et al.[3] under the reinforcement learning framework, and is of particular interest in neuroscience research. Iigaya et al. consider situations in which the decision-maker can request advance information about the availability of a probabilistic reward, and propose that decision-makers are motivated to obtain information because they derive positive (negative) utility from the positive (negative) reward prediction errors produced by the information. Importantly, the model considers situations in which the information refers to a single utility-relevant outcome and its utility depends on the (recursive) total value of that outcome. Because of this assumption, their model predicts that the decision-makers' inquiries would be invariably directed to the high-variance lottery in our experimental paradigm, as information about this lottery produces the largest reward prediction errors, and hence the strongest anticipatory utility (Supplementary Note 3).

The assumption of recursivity adopted by Iigaya et. al. is common with Kreps and Porteus's model[2] and is attractive on computational grounds, as it ensures tractability and the temporal consistency of predicted choice behaviour.

However, the pervasive influence of ΔEV evident in our data suggests that this assumption is too restrictive to account for curiosity. To explain the influence of ΔEV in our task, we postulated that, in addition to being curious about the total reward, decision-makers may independently savour information about individual components of that reward. Moreover, the degree of savouring increases nonlinearly (as a convex function) with the anticipated probability of the outcome, so that rewards that are more certain to be obtained are savoured more than those that are more uncertain, by a factor greater than the increase in probability itself (Supplementary Note 2). Because of this assumption, the total utility from savouring does not necessarily depend simply on the distribution of the total reward from all sources that will be received, as is generally assumed in recursive models. This minimal modification of previous models suffices to explain the effects of ΔEV in our data.

Our model makes novel predictions that can be tested in future investigations. First, future empirical studies can provide additional tests of our prediction that information demand has a nonlinear sensitivity to probability (rather than depending only on the EV of a possibly nonlinear function of the final total reward; Supplementary Note 2). Second, a particularly interesting question is whether attribute-specific anticipatory utility influences instrumental information demand. The need to investigate complex situations is not restricted to curiosity, but is characteristic of many instrumental decisions (for example, an investor handling a multi-asset portfolio or a consumer shopping for a car). Emerging evidence suggests that people are not always optimal in their strategies for gathering instrumental information, but have attentional and learning biases towards irrelevant items[31,32] or fail to acquire the information that provides the most efficient reduction in uncertainty[33–35]. It is of considerable interest to determine the extent to which these inefficiencies may be explained by attribute-specific anticipatory utility.

While our participants' anticipatory utility was correlated in the domain of gains and losses, our results leave open the possibility that savouring and dread have important dissociations. Whereas savouring was highly robust in several groups of participants, dread was weaker and more sensitive to context. This is consistent with previous findings suggesting that dread shows higher variability. For instance, humans have been reported to attend preferentially to more salient (worse) outcomes[20]—a tendency opposite to the one we find—but also to avoid medical information in proportion to its potential seriousness[19], consistent with the present results. Moreover, dread was a notable exception among the parameters we measured in that it showed a weak association with personality measures of the tendency to approach rewards and risk-avoidant attitudes. Thus, the mechanisms of dread and their dependence on context and personality traits may be important questions for future investigations.

From a mechanistic perspective, our results imply that mental activity (that is, attention, memory or belief updating) is recruited by pathways that signal value or uncertainty independent of instrumental incentives. Neural investigations have identified correlates of value and uncertainty associated with both instrumental[36,37] and non-instrumental valuation[38–41], and it will be important to understand better how these systems are functionally related[42]. Recent studies have shown that exploration becomes more sophisticated with age[43,44], raising the interesting question of whether a similar finding applies to anticipatory utility, that is, whether younger individuals show a stronger influence of (potentially simpler) attribute-specific anticipatory utility while the (potentially more sophisticated) interest in total utility becomes stronger with age.

## Methods

**Participants.** We collected data from 298 participants, of whom 129 were recruited from the Columbia University community and tested in two cohorts in the laboratory (laboratory I: individual testing; $n=40$; laboratory II: group testing; $n=89$). The remaining 169 performed the task on the online platform Amazon Turk. Laboratory I included 22 women and 18 men. Laboratory II included 47 women and 42 men whose ages were in the range 18–48 years (median: 22 years; mean: 23.0 years). Information on the gender of Amazon Turk participants and the age of laboratory I and Amazon Turk participants was not collected. Task instructions and contingencies were programmed in MATLAB PsychToolbox with a 24-inch monitor for laboratory participants, and on Python/Amazon psiTurk for Amazon Turk participants. All participants provided informed consent. All of the procedures were approved by the Institutional Review Board of Columbia University.

**Experimental design.** All participants completed a single testing session divided into blocks of 90 trials, with each block representing 1 task. Our focus was on the two-lottery observing task, which was run in two versions to test preferences for information about either monetary gains or monetary losses. In addition, participants completed two control tasks designed to test their willingness to pay for information (WTP task) and their risk sensitivity (risk task). Finally, participants in laboratory II also completed personality questionnaires.

*Two-lottery observing task, gain domain.* On each trial, participants were presented with two lotteries that differed in their EV and variance (the range of points they provided), and were instructed that: (1) each lottery could deliver five discrete amounts evenly distributed across its range; (2) the computer will randomly draw one of the available amounts from each lottery; and (3) it will calculate the trial's payoff as the sum of the draws. While both prizes were relevant to the trial's payoff, participants were asked to choose one prize whose value they wished to reveal immediately, while remaining ignorant about the prize from the remaining draw. The participants' choice did not affect the trial's payoff, which was determined strictly by the sum of the random draws.

The two lotteries presented on each trial were depicted visually, as shown in Fig. 1a, by means of two dark grey bars whose midpoint indicated the lottery EV (also marked by a line and numerical value in blue), and whose length indicated the lottery range. The dark bars were superimposed on scale bars with a lighter background—two rectangles that were positioned symmetrically to the right and left of the screen centre, and whose length indicated a constant range (0–500 points). On declaring their observing decision, participants were shown the precise prize that had been drawn from the chosen lottery by means of a red bar and a numerical value written in white letters above the screen centre. The prize from the non-chosen lottery was not revealed, and a question mark was displayed to emphasize the fact that this prize was nevertheless added to the earnings.

On each trial, one lottery had a high variance and the other had a low variance equal to, respectively, 1,800 and 200 points in laboratory I (corresponding to ranges of 120 and 40 points), and 3,200 and 800 points in laboratory II and Amazon Turk (corresponding to ranges of 160 and 80 points).

In addition, the relative EV of the two lotteries (ΔEV, the EV of the high-variance lottery minus the EV of the low-variance lottery) was drawn randomly with uniform probability from a set of 9 possible values (−90, −70, −20, −5, 0, 5, 20, 70 and 90 points). A random value was then added to both lotteries, jittering the total EVs without altering the relative variance or ΔEV. Thus, the EVs of the high-variance lotteries were uniformly distributed between 101–399 points in laboratory I and 171–330 points in laboratory II and Amazon Turk. The locations of high-variance and low-variance lotteries (left or right) were randomized across trials.

*Two-lottery observing task, loss domain.* This experiment was identical to that for the gain domain, except that participants received an endowment of 1,500 points at the beginning of the block and were instructed that this endowment would be reduced by the sum of random draws. The sequence of events in a trial, number of trials per block, method of calculating the bonus, and variance and EV of the lotteries were identical to those in the gain domain.

Importantly, the visual depiction of the lotteries was also identical (Fig. 1a). Whereas in the gain domain, a higher location on the screen indicated a better outcome (higher possible gain), it indicated a worse outcome (higher possible loss) in the loss domain, allowing us to avoid possible confounds related to upper/lower visual field preference. To clarify to participants that the lotteries yielded not gains but losses in this task, the displayed numbers (lotteries' EVs, scale labels and revealed prizes) were preceded by a minus sign (−).

*WTP task.* The WTP task tested the extent to which participants were willing to pay for advance information about a monetary outcome. Participants were given two lotteries on each trial and selected a single lottery from which they wished their prize to be drawn. Therefore, in contrast with the observing task, in the WTP task, the participants' payoffs were contingent on their choices. The two lotteries had equal variance (of 3,200 points), but differed independently in their EV and the availability of information. One lottery provided immediate information about the precise prize that had been drawn, while the other did not, and informativeness was signalled by lottery colour (green or purple, counterbalanced across participants). By manipulating information availability independent of ΔEV, we could estimate the price participants were willing to pay to obtain information.

The visual displays and procedures were similar to those in the two-lottery task. Values of ΔEV were drawn from the set of −40, −20, −5, −2, 0, 2, 5, 20 and 40 points—a smaller range relative to the two-lottery task to allow for the possibility that the willingness to pay would be small[45]. As in the two-lottery paradigm, EVs of respective lotteries were jittered across trials and the EV of the informative lottery followed a uniform distribution ranging between 121 and 380 points.

*The risk task (gain and loss domains).* This conventional risk-taking paradigm was used to verify that participants responded to incentives. Participants received a risky option and a safe option on each trial, and were asked to select from which they wished to receive a payoff. The risky lottery had a variance of 1,800 points in laboratories I and II and 3,200 on Amazon Turk (corresponding to ranges of 120 and 160 points, respectively) and the safe option had a single possible outcome (0 variance and range). As in the WTP task, participants received a payoff equal to a single random draw from the lottery of their choice. To avoid confounding risk and informational preferences, participants did not receive immediate feedback about the outcome of each trial. The visual displays and procedures, as well as the levels of EV and ΔEV (the EV of the risky lottery minus the safe value), were identical to those on the corresponding two-lottery observing tasks.

*Personality questionnaires.* For participants in laboratory II, we explored the association between the information-seeking behaviour and personality traits by collecting their responses to personality questionnaires (see the main text for adopted measures), as well as their age, gender, the degree they were pursuing and the field of their major (all laboratory II participants were Columbia University students). This information was collected after the participants completed the behavioural tasks and took between 15 and 30 min to complete.

**General procedures, instructions and payment.** Of the 298 participants who completed the two-lottery observing task, 139 completed it only in the gain domain (40 in laboratory I and 99 on Amazon Turk). These participants also completed the WTP and risk tasks in the gain domain. The remaining 159 participants completed the observing task in both the gain and loss domains (70 on Amazon Turk and 89 in laboratory II) and also completed the risk task in both domains. We deployed tasks in the different orders across groups, such that we could verify that our effects were not caused by the order of testing (Supplementary Table 1). Data collection and analysis were not performed blind to the conditions of the experiments.

The tasks were administered in trial blocks separated by a brief pause. For each block, participants were instructed to hold their right hand steady on the keyboard with the index and middle fingers positioned, respectively, above the left and right arrow keys and, after the onset of the lottery display, to indicate which lottery they wished to play or observe by pressing the corresponding key (left or right). There was no time pressure for making the selection. The participant's response was followed by a feedback screen displaying the requested information (either the value of the realized prize or a '?' depending on the participant's choices and the task), a second screen indicating the trial progression (for example, '12/90 trials') and the onset of the following trial. The feedback and trial progression screens lasted 1 s for all trial types in all tasks.

*Instructions and payment.* The experimenter provided each participant with a general explanation at the start of a session and additional task-specific instructions at the start of each trial block.

Participants were informed that they would receive a base payment for completing the experiment (US$12 h$^{-1}$ for laboratory I and laboratory II; US$1 for Amazon Turk), as well as a bonus at the end of each block, which was equal to the outcome of one trial randomly selected from those that had been played in that block. Participants were instructed that the dollar amount of the bonus would be proportional to the point value of the selected trials (500 points were worth US$1; the participants who completed the observing task in the loss domain were specifically informed of the conversion rate in the instructions). At the end of each block, participants saw a display screen reporting the dollar values of the bonus on the current block and the total bonus earned so far in the session.

An important concern was that participants may incorrectly understand the instructions—specifically, the fact that their choices influenced their payoffs only in the WTP and risk tasks, but not in the observing paradigms. To prevent this possibility, during the instruction phase we made it very clear to the participants that their outcomes on the observing task would be determined purely by chance, and that they would receive the prizes drawn from both lotteries regardless of their observing decision. Furthermore, for laboratory participants, we confirmed that they understood the instructions during debriefing at the end of the session (laboratory I) or by comprehension quizzes before the task (laboratory II). Since we could not deliver the instructions in person for Amazon Turk participants, we tried to minimize the possibility of confusion between tasks and domains by presenting the more customary risk task first and the observing tasks later in the session, and the loss domain tasks (including the observing and risk task) were conducted after those for gains. The task order was reversed for the laboratory II participants, allowing us to verify that our population-level inference does not depend on these specific task orders (Supplementary Table 1).

*Sample sizes.* We had little a priori basis for determining the sample size. Therefore, we chose to test a relatively large number of participants ($n=298$) to allow for potential individual variation. In addition, since an important part of our analysis was at the individual level, we collected a generous number of trials from each participant (90 trials for fitting two-parameter psychometric curves). All of the sample sizes were predetermined and not altered based on the results.

**Data analysis.** *Pre-processing.* We discarded the data from participants who did not respond to monetary incentives (that is, those who were insensitive to ΔEV in either the WTP task or the risk task (formally, these were individuals who had a $w_{\Delta EV}$ parameter that was not significantly greater than 0 ($P>0.05$); see below, equation (1))). Based on these criteria, we discarded the data from 22 of the 139 participants who completed the observing task in the gain domains (1/40 in laboratory I and 21/99 on Amazon Turk) and 19 of the 159 participants who completed it for both gains and losses (16/70 on Amazon Turk and 3/89 in laboratory II). Thus, the analysis focuses on 257 participants who completed the observing task in the gain domain (39 in laboratory I, 86 in laboratory II and 132 on Amazon Turk), of whom 140 also performed the task in the loss domain (86 in laboratory II and 54 on Amazon Turk).

Note that these exclusion criteria are independent of our inference on participants' behaviour in the observing task. The rationale behind this exclusion is that, if participants did not show sensitivity to monetary incentives, this probably reflected inattention to the task, making it difficult to interpret their observing behaviour. Indeed, we had to exclude a sizable proportion of Amazon Turk participants, suggesting that these participants were motivated to finish the task quickly without paying attention to instructions.

*Choice modelling.* We used maximum-likelihood estimation to fit individual participants' choices with a two-parameter logit model:

$$P\,(\text{choice}) = \frac{1}{1+\exp\left(-\left(w_0 + w_{\Delta EV}\cdot \Delta EV\right)\right)} \tag{1}$$

in which $w_0$ and $w_{\Delta EV}$ are free (estimated) parameters, $P$(choice) is the probability of choosing one of the options, and ΔEV is the difference between the EVs of the two lotteries (standardized to the range of −1 to 1 for the parameter estimation).

In the observing task, $P$(choice) was defined as the probability of observing the high-variance lottery, and ΔEV as the EV of the high-variance lottery minus the EV of the low-variance lottery. As described in Supplementary Note 1, the parameter $w_{\Delta EV}$ (the slope of the psychometric function; Fig. 3) indexes the propensity to observe based on attribute-specific anticipatory utility, with positive values indicating preferences for advance information about the more desirable outcome. More precisely, as explained in Supplementary Note 2, a positive value indicates that positive rewards that one is certain to receive are savoured more than those that remain mere possibilities. In contrast, $w_0$ (the vertical shift of the psychometric function) indexes the propensity to observe based on uncertainty reduction, with positive and negative values indicating, respectively, a preference for the early or late resolution of uncertainty. Thus, for the observing paradigm, we refer to $w_0$ as $w_{var}$.

For the WTP task, we modelled $P$(choice) as the probability of choosing the informative lottery, and ΔEV as the EV of the informative lottery minus the EV of the uninformative lottery. Therefore, $w_{\Delta EV}$ indexes the propensity to choose the lottery with higher EV, and $w_0$ indexes the propensity to choose the informative lottery. For the risk task, $P$(choice) was the probability of choosing the risky lottery, and ΔEV was the EV of the risky lottery minus the value of the safe option. Therefore, $w_{\Delta EV}$ indexes the propensity to choose the option with higher EV or safe value, and $w_0$ indexes the propensity of risk taking.

Estimating the two parameters in equation (1) amounts to conventional logistic regression. Preliminary analyses showed that ridge logistic regression improved the model fit compared with unregularized logistic regression, as evaluated by within-participant cross-validation, provided that the regularization term $\lambda$ was in the range 0.001–0.1, with little difference within this range. The parameter estimates reported in the main text are under regularization with $\lambda=0.01$.

To evaluate the statistical significance of individual parameters, we constructed null-hypothesis distributions by randomly shuffling the trial labels for 1,000 iterations, separately for $w_0$ and $w_{\Delta EV}$ (for $w_0$, the variance labels were shuffled within trials while maintaining absolute values of ΔEV; for $w_{\Delta EV}$, the ΔEV labels were shuffled across trials). Individual parameters were evaluated against these null distributions (two-sided test, $\alpha=0.05$).

Population-level statistical inference used a two-sided Wilcoxon signed rank test unless otherwise noted. We report $Z$ statistics and the corresponding effect size measure, $r$[46]. The 95% CI of the effect size $r$ was obtained by non-parametric bootstrap (random sampling of participants with replacement; 1,000 iterations) and the bias-corrected and accelerated method[47].

The PSE, indicating the ΔEV at which the participant was indifferent between the two alternatives, was calculated based on the fitted logistic functions (equation (1)). The 95% CI of the PSE was obtained by non-parametric bootstrap (random resampling of trials with replacement; 1,000 iterations) and the bias-corrected and accelerated method[47].

*Reaction time analysis.* Trials in two-lottery observing tasks in which the reaction time was shorter than 0.2 s or longer than 5 s, or was more than 2 standard deviations above each participant's mean, were discarded and not analysed further. The remaining values were *Z* scored for each participant and aggregated across participants.

To examine the relationship between choice preferences and reaction time, we adopted a prediction from the standard drift diffusion model (DDM) given the choice parameters estimated in equation (1)[48]. The DDM predicts a quantitative relationship between reaction time and choice difficulty. Specifically, the reaction time for different levels of ΔEV is predicted to follow $\tau_i = \tanh\left(\frac{\mathrm{DV}_i}{2}\right)/\mathrm{DV}_i$, where $\mathrm{DV}_i$ is the decision variable used in equation (1), $\mathrm{DV}_i = w_{\mathrm{var}} + w_{\Delta\mathrm{EV}} \cdot \Delta\mathrm{EV}_i$, and $i$ indexes the level of ΔEV. To test this prediction, for each subgroup classified based on the significance and sign of the parameters of individual observing behaviours, we first estimated the group-level parameters $w_{\mathrm{var}}$ and $w_{\Delta\mathrm{EV}}$ from the choices averaged over participants using the same procedure as the individual-level model fit. We then fit the normalized reaction time as a linear function of $\tau_i$, namely:

$$\mathrm{RT}_{ij} = A\tau_i + t_0 + \varepsilon_{ij} \quad (2)$$

where $j$ indexes the individual participant and $A$ indexes the extent to which reaction times were sensitive to choice difficulty. The free parameters $A$ and $t_0$ were estimated by least-squares regression, and the significance of $A$ was evaluated under the assumption of asymptotic normal distribution. Note that we adopted this approach merely to conveniently capture the relationship between ΔEV and reaction time; we do not claim a mechanistic account of decision-making described by the DDM.

*Personality questionnaires.* Linear regression modelling was conducted to examine whether personality measures would predict parameter estimates of individual observing behaviour ($w_{\mathrm{var}}$ and $w_{\Delta\mathrm{EV}}$, in gain and loss domains, respectively). Since we have a relatively large number of predictors (nine personality questionnaires and four demographic variables, namely age, gender (female versus male; all participants self-reported their gender as one of these two), education (undergraduate versus graduate students) and field (coded as quantitative versus non-quantitative by an author))), we deployed the automatic feature selection procedure using Lasso regularization (the lasso function in MATLAB), which determined the regularization level by tenfold cross-validation. At the automatically selected levels of regularization, for $w_{\Delta\mathrm{EV}}$ in the gain domain and $w_{\mathrm{var}}$ (in both gain and loss domains), none of the predictors survived. For $w_{\Delta\mathrm{EV}}$ in the loss domain, BIS, BAS, DOSPERT, gender and education were selected (Supplementary Table 2).

Since we found that $w_{\Delta\mathrm{EV}}$ in the loss domain failed to reach significance at the population level among participants who completed the observing task in the loss domain before the gain domain (Supplementary Table 1; see main text), we examined the interaction between the personality measures selected above and the task order in an additional linear regression (without regularization). None of the interaction terms was significant, confirming that the original regression was not confounded by the task order (Supplementary Table 3).

Lastly, to explore which submeasures of BAS and DOSPERT drive the association, we ran another linear regression, replacing the BAS and DOSPERT main scores with their submeasure scores, without regularization (Supplementary Table 4).

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Requests for the data can be sent via email to the corresponding author.

## Code availability

Requests for the code used for all analyses can be sent via email to the corresponding author.

Received: 2 April 2018; Accepted: 12 March 2019;
Published online: 15 April 2019

## References

1. Berlyne, D. E. A theory of human curiosity. *Br. J. Psychol.* **45**, 180–191 (1954).
2. Kreps, D. M. & Porteus, E. L. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica* **46**, 185–200 (1978).
3. Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J. & Dayan, P. The modulation of savouring by prediction error and its effects on choice. *eLife* **5**, e13747 (2016).
4. Loewenstein, G. The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* **116**, 75–98 (1994).
5. Gottlieb, J. & Oudeyer, P. Y. Toward a neuroscience of active sampling and curiosity. *Nat. Rev. Neurosci.* **19**, 758–770 (2018).
6. Gottlieb, J., Oudeyer, P. Y., Lopes, M. & Baranes, A. Information seeking, curiosity and attention: computational and empirical mechanisms. *Trends Cogn. Sci.* **17**, 585–593 (2013).
7. Kidd, C. & Hayden, B. Y. The psychology and neuroscience of curiosity. *Neuron* **88**, 449–460 (2015).
8. Loewenstein, G. & Molnar, A. The renaissance of belief-based utility in economics. *Nat. Hum. Behav.* **2**, 166–167 (2018).
9. Vasconcelos, M., Monteiro, T. & Kacelnik, A. Irrational choice and the value of information. *Sci. Rep.* **5**, 13874 (2015).
10. Eliaz, K. & Schotter, A. Experimental testing of intrinsic preferences for noninstrumental information. *Am. Econ. Rev.* **97**, 166–169 (2007).
11. Zentall, T. R. & Stagner, J. P. Do pigeons prefer information in the absence of differential reinforcement? *Learn. Behav.* **40**, 465–475 (2012).
12. Bromberg-Martin, E. S. & Hikosaka, O. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* **63**, 119–126 (2009).
13. Blanchard, T. C., Hayden, B. Y. & Bromberg-Martin, E. S. Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* **85**, 602–614 (2015).
14. Baranes, A. F., Oudeyer, P. Y. & Gottlieb, J. Eye movements encode epistemic curiosity in human observers. *Vis. Res.* **117**, 81–90 (2015).
15. Daddaoua, N., Lopes, M. & Gottlieb, J. Intrinsically motivated oculomotor exploration guided by uncertainty reduction and conditioned reinforcement in non-human primates. *Sci. Rep.* **6**, 20202 (2016).
16. Gruber, M. J., Gelman, B. D. & Ranganath, C. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84**, 486–496 (2014).
17. Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil. Trans. R. Soc. Lond. B* **362**, 933–942 (2007).
18. Loewenstein, G. Anticipation and the valuation of delayed consumption. *Econ. J.* **97**, 666–684 (1987).
19. Caplin, A. & Leahy, J. Psychological expected utility theory and anticipatory feelings. *Q. J. Econ.* **116**, 55–79 (2001).
20. Kahnt, T., Park, S. Q., Haynes, J. D. & Tobler, P. N. Disentangling neural representations of value and salience in the human brain. *Proc. Natl Acad. Sci. USA* **111**, 5000–5005 (2014).
21. Carver, C. S. & White, T. L. *Behavioral Avoidance/Inhibition (BIS/BAS) Scales* (Measurement Instrument Database for the Social Sciences, 2013).
22. Gard, D. E., Gard, M. G., Kring, A. M. & John, O. P. Anticipatory and consummatory components of the experience of pleasure: a scale development study. *J. Res. Pers.* **50**, 1086–1102 (2006).
23. Foa, E. B. et al. The obsessive–compulsive inventory: development and validation of a short version. *Psychol. Assess.* **14**, 485–496 (2002).
24. Spielberger, C. D. *Manual for the State-Trait Anxiety Inventory* (Mind Garden, 1983).
25. Radloff, L. S. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psych. Meas.* **1**, 385–401 (1977).
26. Litman, J. A. Interest and deprivation factors of epistemic curiosity. *Pers. Individ. Dif.* **44**, 1585–1595 (2008).
27. Litman, J. A. & Spielberger, C. D. Measuring epistemic curiosity and its diversive and specific components. *J. Pers. Assess.* **80**, 75–86 (2003).
28. Blais, A.-R. & Weber, E. U. A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgm. Decis. Mak.* **1**, 33–47 (2006).
29. Van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R. & de Lange, F. P. Induction and relief of curiosity elicit parietal and frontal activity. *J. Neurosci.* **38**, 2579–2588 (2018).
30. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P. & Pezzulo, G. Active inference: a process theory. *Neural Comput.* **29**, 1–49 (2017).
31. Anderson, B. The attention habit: how reward learning shapes attentional selection. *Ann. NY Acad. Sci.* **1369**, 24–39 (2016).
32. Sharot, T. The optimism bias. *Curr. Biol.* **21**, 941–945 (2011).
33. Barbaro, L., Peelen, M. V. & Hickey, C. Valence, not utility, underlies reward-driven prioritization in human vision. *J. Neurosci.* **37**, 10438–10450 (2017).
34. Morvan, C. & Maloney, L. Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Comput. Biol.* **8**, e1002342 (2012).
35. Hunt, L. T., Rutledge, R. B., Malalasekera, W. M., Kennerley, S. W. & Dolan, R. J. Approach-induced biases in human information sampling. *PLoS Biol.* **14**, e2000638 (2016).
36. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363–375 (2005).

37. Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
38. Schultz, W. et al. Explicit neural signals reflecting reward uncertainty. *Phil. Trans. R. Soc. Lond. B* **363**, 3801–3811 (2008).
39. O'Neill, M. & Schultz, W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* **68**, 789–800 (2010).
40. Monosov, I. E. & Hikosaka, O. Selective and graded coding of reward uncertainty by neurons in the primate anterodorsal septal region. *Nat. Neurosci.* **16**, 756–762 (2013).
41. Monosov, I. E., Leopold, D. A. & Hikosaka, O. Neurons in the primate medial basal forebrain signal combined information about reward uncertainty, value, and punishment anticipation. *J. Neurosci.* **35**, 7443–7459 (2015).
42. Levy, I., Lazzaro, S. C., Rutledge, R. B. & Glimcher, P. W. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* **31**, 118–125 (2011).
43. Zajkowski, W. K., Kossut, M. & Wilson, R. C. A causal role for right frontopolar cortex in directed, but not random, exploration. *eLife* **6**, e27430 (2017).
44. Wilson, R. C. & Niv, Y. Inferring relevance in a changing world. *Front. Hum. Neurosci.* **5**, 189 (2011).
45. Bennett, D., Bode, S., Brydevall, M., Warren, H. & Murawski, C. Intrinsic valuation of information in decision making under uncertainty. *PLoS Comp. Biol.* **12**, e1005020 (2016).
46. Fritz, C. O., Morris, P. E. & Richler, J. J. Effect size estimates: current use, calculations, and interpretation. *J. Exp. Psychol. Gen.* **141**, 2–18 (2011).
47. DiCiccio, T. J. & Efron, B. Boostrap confidence intervals. *Stat. Sci* **11**, 189–228 (1996).
48. Shadlen, M. N., Hanks, T. D., Churchland, A. K., Kiani, R. & Yang, T. The speed and accuracy of a simple perceptual decision: a mathematical primer. in *Bayesian Brain: Probabilistic Approaches to Neural Coding* (eds Doya, K. et al.) Ch. 10 (MIT Press, 2006).


## Acknowledgements

The work was supported by a Human Frontiers Cross-Disciplinary Fellowship (to A.B.), Presidential Scholars in Science and Society Seed Grant at Columbia University (to M.W. and J.G.), Research Initiatives in Science and Engineering Seed Grant at Columbia University (to M.W. and J.G.), National Science Foundation grant SES-1426168 (to M.W.) and the Cognitive and Behavioral Economics Initiative at Columbia University (to S.R.). We thank M. Jameson and J. Capaldi for expert administrative assistance. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

J.G. designed the experiment. K.K., S.R. and A.B. implemented the task and collected the data. K.K. analysed the data. M.W. wrote the computational model. K.K., S.R., M.W. and J.G. interpreted the results and wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-019-0589-3.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to K.K.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Kenji Kobayashi

Last updated by author(s): Jan 28, 2019

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size ($n$) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. $F$, $t$, $r$) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's $d$, Pearson's $r$), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Custom codes (MATLAB R2014b, Psychtoolbox, Javascript), Amazon PsiTurk |
| Data analysis | Custom codes (MATLAB R2014b) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings are available from the corresponding author upon request.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Study description | Quantitative experimental. |
| Research sample | 129 Columbia University students (69 female, 60 male) & 169 on-line web-survey participants on Psiturk platform (no demographic information available). |
| Sampling strategy | Lab participants were recruited via fliers ("Lab I") or through participant pool of Columbia University Department of Economics ("Lab II"); on-line participants were recruited on PsiTurk platform ("AT"). Because our study is, to the best of our knowledge, the first one examining our empirical questions, we had little a priori basis for determining the sample size. We chose to test a relatively large number of participants (n = 298) in order to allow for potential individual variation. The sample size was pre-determined and not altered based on the results. |
| Data collection | Computer-based experiment in lab booths & on-line web survey on participants' own web browser. |
| Timing | Lab I: April to June 2015, Lab II: September 2018, AT: July to September 2015 (tested on "Gain domain" only) and February to March 2016 (tested both on "Gain domain" and "Loss domain"). |
| Data exclusions | Participants who did not respond to monetary incentives were discarded, as described in Methods section. This exclusion procedure does not use the data that support our main findings and is independent of our hypothesis-testing analyses. The exclusion was deployed before our main analyses. The results of analyses that include those participants are also reported in Supplementary Table 1. |
| Non-participation | No lab participants dropped out or declined during the experiment. Some AT participants may or may not have dropped out, but they were never included in our dataset. |
| Randomization | Participants in Lab I and AT were not assigned to groups. We conducted two versions of AT experiment, separated in time (see above). Participants in Lab II were randomly assigned to two groups, which differed in the order of task deployment (Methods). |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | Antibodies |
| ☒ | ☐ | Eukaryotic cell lines |
| ☒ | ☐ | Palaeontology |
| ☒ | ☐ | Animals and other organisms |
| ☐ | ☒ | Human research participants |
| ☒ | ☐ | Clinical data |

## Methods

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | ChIP-seq |
| ☒ | ☐ | Flow cytometry |
| ☒ | ☐ | MRI-based neuroimaging |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | See above. |
| Recruitment | Lab participants were recruited via fliers ("Lab I") or through participant pool of Columbia University Department of Economics ("Lab II"); AT participants were recruited on PsiTurk platform. These recruitment procedures may have introduced selection bias; those who signed up for experiments may be more likely interested in psychological / economic experiments (Lab) or earning compensation as quickly and efficiently as possible (AT). Such a bias, however, is not predicted to yield our results. |
| Ethics oversight | The Institutional Review Board of Columbia University |

Note that full information on the approval of the study protocol must also be provided in the manuscript.