bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Title: Hunger for Knowledge: How the Irresistible Lure of Curiosity is Generated in the Brain

Authors: Johnny King L Lau[1,2]*, Hiroki Ozono[3], Kei Kuratomi[4], Asuka Komiya[5], Kou Murayama[1,2*]

Affiliations:

[1]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK.

[2]Centre for Integrative Neuroscience and Neurodynamics, University of Reading, Reading, UK.

[3]Faculty of Law, Economics and Humanities, Kagoshima University, Kagoshima, Japan.

[4]Faculty of Psychology, Aichi Shukutoku University, Nagakute-city, Japan.

[5]Graduate School of Integrated Arts and Sciences, Hiroshima University, Higashi-Hiroshima city, Japan.

*Correspondence to: johnny.lau@reading.ac.uk (Johnny Lau), k.murayama@reading.ac.uk (Kou Murayama)

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introductory Paragraph

Curiosity is often portrayed as a desirable feature of human faculty. For example, a meta-analysis revealed that curiosity predicts academic performance above and beyond intelligence[1], corroborating findings that curiosity supported long-term consolidation of learning[2,3]. However, curiosity may come at a cost of strong seductive power that sometimes puts people in a harmful situation. Here, with a set of three behavioural and two neuroimaging experiments including novel stimuli that strongly trigger curiosity (i.e. magic tricks), we examined the psychological and neural mechanisms underlying the irresistible lure of curiosity. We consistently demonstrated that across different samples people were indeed willing to gamble to expose themselves to physical risks (i.e. electric shocks) in order to satisfy their curiosity for trivial knowledge that carries no apparent instrumental values. Also, underlying this seductive power of curiosity is its incentive salience properties, which share common neural mechanisms with extrinsic incentives (i.e. hunger for foods). In particular, the two independent fMRI experiments using different kinds of curiosity-stimulating stimuli found replicable results that acceptance (compared to rejection) of curiosity/incentive-driven gambles was accompanied with an enhanced activity in the striatum.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Curiosity is a fundamental part of human motivation that supports an enormous variety of human intellectual behaviours, ranging from early learning in children to scientific discovery [4–6]. The critical importance of curiosity in human intellectual behaviour is succinctly expressed in Albert Einstein's famous quote "*I have no special talent. I am only passionately curious.*" Moreover, empirical literature has revealed a number of positive outcomes associated with curiosity over the life span [1,7–9].

However, in both historic and modern literature, the positive portrayal of curiosity is often compromised by its inherent negative aspect: strong seductive power [10]. In Greek mythology, for example, after losing his beloved *Eurydice* to the underworld, *Orpheus* convinced the gods to let him take her back to the world of the living on the condition that he would not look back until they had returned. *Orpheus* could not help but looked back; he succumbed to curiosity and lost *Eurydice*. This theme, which appears repeatedly in classic ancient anecdotes (e.g., *Pandora*, *Psyche*, *Eve*), illustrates the irresistible power of curiosity, which biases our decision-making despite the knowledge of consequential negative outcomes. Previous literature has indicated that both animals [11–13] and humans [3,14–16] are willing to pay small amounts of costs to satisfy curiosity for the knowledge about a future reward that cannot be changed. For example, people are willing to sacrifice parts of future monetary reward in exchange for immediate information about the outcome of a monetary lottery, even though that information cannot be used to alter lottery outcomes[14]. However, little research has directly addressed the nature of the strong lure of curiosity for non-instrumental trivial facts, which may overcome even significant risks.

Herein, we report a set of behavioural and neuroimaging experiments, showing that i) people are willing to expose themselves to physical risks (i.e. electric shocks) to satisfy their

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

curiosity for trivial, non-consequential knowledge (i.e. magic tricks, trivia), and ii) the seductive power of curiosity shares the same motivational processes as extrinsic incentives (i.e. foods).

In the initial behavioural experiment, participants (N=17) were presented with short (8 to 46s; mean = 22.22s; median = 20s) video clips of magic tricks performed by professional magicians (filmed specifically for this study; Supplementary videos 1-2), followed by a wheel of fortune visually depicting the probability of winning (versus losing) a lottery in each trial (Fig. 1a). Participants were then asked to make a decision on whether they would gamble to take the lottery or not. They were told that, if they accept the lottery and win, they can see the secret behind the magic trick after the experiment (with a certain probability; 16.7% - 83.3%; see Methods). The experiment also included food trials, in which participants were presented with food pictures, and they were told that they can eat the food after the experiment if they accept the lottery and win. Critically, for both trials, if participants accepted the lottery and lost, they were told that they would receive electric shock after the experiment. The magnitude of electric shock was calibrated and demonstrated to participants before the experiment. Electric shock was not delivered during the experimental task, as pilot data indicated that expectation of electric shock has more sustained effects on participants' fear perception than receiving actual electric shock multiple times.

[INSERT FIGURE 1]

Generalised linear mixed-effects modelling showed that participants were more likely to reject the lottery as the presented outcome probability of receiving electric shock increases [$\text{Exp}(\beta) = 2.30$ (for curiosity condition) & 2.91 (for food condition), $Ps < 0.01$]. This suggests that electric shock worked effectively as an aversive stimulus. Importantly, curiosity ratings of magic tricks and desirability ratings of food items both positively predicted the "accept" decision

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

above and beyond the outcome probability [Exp (β) = 2.32 (for magic curiosity) & 2.94 (for food desirability), *Ps* < 0.01]. In other words, curiosity for non-consequential knowledge (i.e. magic solution), as well as incentive value of foods, can bias decision-making entailing physical risks (Supplementary Fig. 1 and Supplementary Table 1; also see Fig. 1b for results integrated with following fMRI experiments).

How does curiosity bias our risky decision-making? There are two possibilities. The first possibility is that curiosity distorts the perception of probability of risks (i.e. when people are curious and in a positive mood, they overestimate the probability of winning the lottery; [17]). In a follow-up behavioural experiment, participants (N=29) rated their subjective chance of winning when presented with a lottery, after seeing a curiosity/hunger evoking stimulus (see Supplementary Methods). However, curiosity rating for magic tricks [Exp (β) = 1.03, *P* = 0.13] or desirability rating for foods [Exp (β) = 1.05, *P* = 0.17] did not predict the participant's subjective probability of winning [Bayes Factor (BF_{01}) for Magic = 0.00524, BF_{01} for Food = 0.00658; very strongly supportive of the null hypothesis].

Another possibility is that curiosity triggers *incentive salience* for unknown information[18]. Expectation of reward often increases incentive salience (or "wanting") and this likely promotes irrational and risky behaviour through modulating the neurobiological processes in the reward system [19–21]. Similar psychological mechanisms may underlie the seductive power of curiosity. Indeed, previous functional magnetic resonance imaging (fMRI) research has revealed that the state of curiosity (e.g. when participants are waiting for the answer to a trivia question) activates the striatum[2,3] --- a key brain area for processing incentive salience [22]. To test this hypothesis, we conducted a neuroimaging experiment with fMRI using a similar experimental paradigm as in the initial behavioural investigation. If the seductive power of curiosity is driven by increased

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

incentive salience, the striatum will be activated when participants accept the lottery[23,24]. A 2 (stimulus type: magic tricks vs. foods) X 2 (decision: accept vs. reject) analysis of variance (ANOVA) (N=31) on a priori anatomical striatum region of interest (ROI) showed that the striatum was indeed activated when participants accepted the lottery in comparison to when they rejected it, regardless of the stimulus type ($P_{FWE} < 0.05$; Fig. 2a-c). This effect was robustly observed even after controlling for the "win" probabilities ($P_{FWE} < 0.05$; Supplementary Fig. 2) in the ROI analysis. No significant interaction between stimulus type and decision was observed within the ROI. A whole brain analysis also indicated that the striatum showed extensive activation for accepted trials in comparison to rejected trials among all brain areas (Supplementary Tables 2-3).

[INSERT FIGURE 2]

We also examined whether the observed striatal activation mediates the relationship between curiosity/food desirability ratings and the acceptance of the lottery on a trial-by-trial basis. A multilevel mediation analysis showed that this mediation effect is statistically significant ($P_{mediation} < 0.05$; Fig. 2d), further supporting the idea that the curiosity or desire for food induced by the stimuli prompted people to risk of physical pain by recruiting a common brain network.

To further explore the mechanisms underlying curiosity-driven decision that overcomes potential physical risks, we conducted a beta-series correlation functional connectivity analysis using the striatum as the seed [a spherical ROI with a 6 mm-radius was created around the peak voxel of the striatum shown to be activated for the accept (vs. reject) decision in the main fMRI analysis]. This exploratory analysis found weaker connectivity between the left sensorimotor cortex and left striatum, suggesting a *decoupling* effect (i.e. weaker functional correlation) associated with accepting the lottery (Fig. 3, Supplementary Table 4); however statistical

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

analysis only revealed marginal significance (P_{FWE} = 0.056) after cluster-level familywise error correction. The sensorimotor cortex has been implicated in anticipation of aversive signals including pain[25,26]. Our results may suggest that, when the participant's curiosity to see the secret of a magic trick overcame the risk of receiving electric shock, there tended to be a dissociation of incentive salience and expected fear of physical shock in the brain.

[INSERT FIGURE 3]

Magic tricks evoke people's curiosity through sensory (i.e. visual) input. This type of curiosity can be described as "perceptual curiosity"[27]. To examine the robustness of our findings and investigate whether they can be generalised to the type of curiosity that is evoked by conceptual ambiguities or uncertainties (called "epistemic curiosity" [27]), we conducted a replication fMRI experiment, in which we presented trivia questions (e.g., "What is the only food that never spoils?") selected from a standardised database [28] (instead of magic tricks) to stimulate curiosity within an independent sample. A 2 (stimulus type: trivia questions vs. foods) X 2 (decision: accept vs. reject) ANOVA (N=30) was performed to compare the average beta contrast values within the striatum ROI, created based on the activation map in the first fMRI experiment. In the replication fMRI experiment, striatum was again more activated, regardless of the stimulus type, when participants accepted the lottery to satisfy their curiosity in comparison to when they rejected it ($P < 0.05$, see Supplementary Fig. 3a). Moreover, the mediation effect was observed ($P < 0.05$, Supplementary Fig. 3b), and the sensorimotor cortex (based on a functional mask created from the connectivity analysis in the first fMRI experiment) also exhibited functional decoupling with striatal activation ($P < 0.05$, Supplementary Fig. 4). Finally, to explore other brain areas that are associated with a participant's decision-making process, a whole-brain conjunction analysis was performed to examine the overlapping brain areas between

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

the two fMRI samples for the "accept" (vs. "reject") decision. In additional to the caudate nucleus, the dorsal medial prefrontal cortex, premotor cortex and thalamus ($P < 0.00001$, uncor; Fig. 4 & Supplementary Table 2) all exhibited involvement in the decision-making process.

[FIGURE 4]

Incentive salience is a motivational "hot" feeling, potentially prompting people toward impulsive behaviours[29]. The effect of incentive salience would be stronger when people make immediate decisions involving personal dilemmas in contrast to when they make predictions about the behaviours of others with whom they are emotionally detached[30]. To test this possibility, we conducted another experiment in which participants made a prediction about the decision of some hypothetical participants who were depicted to be the subjects to receive actual experimental treatments. In this "prediction" experiment (N = 67), ratings of curiosity and food desirability were still positively associated with the (predicted) decision to accept lottery, but the effect of ratings was smaller when compared with the other experiments requiring actual decision-making ($P<0.05$, for the interaction between Prediction and Rating; see Methods and Supplementary Table 6). These results provide another compelling evidence for the role of incentive salience in curiosity-driven risky decision.

The current study showed that curiosity biases our decision-making by recruiting the same incentive motivation process as extrinsic rewards (e.g. foods). These findings suggest that the seemingly irrational curiosity-based choices that people make can be explained by standard decision-making models, by assuming that knowledge acquisition has inherent (yet rather immediate) rewarding value per se [6,31]. A question remains regarding why incentive salience triggered by curiosity is so strong that it drives people into even self-harming behaviour [32–35]. Also, previous studies have indicated that curiosity is inherently related to the feeling of anxiety

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and fear[10], but our neuroimaging studies did not find clear pattern of activations in the anxiety/fear related areas[36,37]. A cooperative effort from multiple perspectives such as development[6], personality[38], computation[39], and evolution[40], would be required to comprehensively understand the elusive concept of curiosity.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Methods

Participants

The study was approved by the research ethics committees of the University of Reading, UK (ethics approval number: UREC 16/36). Apart from the follow-up 'prediction' experiment, participants were recruited via mailing lists and SONA, a research participation scheme, at the University. Participants provided informed consent, completed and passed a health and safety screening for receiving electric stimulation to confirm that i) they did not have a cardiac pacemaker (or any other devices that can be affected by electric stimulation); ii) they were free from neurophysiological symptoms or conditions including peripheral vascular disease, vasculitis cryoglobulinemia, lupus, tingling or numbness in hands and/or feet. To maximise participants' motivation for food consumption during the experiment, they were required not to eat or drink anything (apart from water) within 2 hours prior to the testing session.

Depending on personal preference, participants were compensated either with course credits at a fixed rate of 1 unit per hour or cash payments for their participation. For behavioural experiments, a fixed rate of £7 per hour was given, while for fMRI experiments the compensation for participation was fixed at £10 per hour. In addition, participants were informed that they may receive extra rewards according to their task performance in the experiment. Each participant took part in only one version of the experiment.

Initial behavioural experiment: Seventeen healthy participants (5 males) with an average age of 23.11 years old (sd= ±5.30) participated.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

fMRI experiment 1 (Curiosity about Magic Tricks): Thirty-two healthy individuals were recruited. One participant accepted every single gamble (100% acceptance on all trials) in the decision-making task and was thus excluded prior to data analysis. The remaining 31 participants (5 males) were on average 19.53 years old (sd= ±1.59) and all right-handed.

fMRI experiment 2 (Curiosity about Trivia): Thirty-two healthy adults took part in the experiment. One of them accepted every single gamble (100% acceptance on all trials). Another participant had pronounced head movements during MRI scan (> 3mm displacement in a motion direction, see Supplementary Methods for further details). Both were excluded prior to data analysis. A sample of the remaining 30 individuals (6 males) had a mean age of 20.3 years old (sd= ±2.49) and all were right-handed.

Follow-up behavioural experiment on estimating the subjective outcome probability (FU1): Twenty-nine healthy participants (7 males) with a mean age of 20.17 years old (sd=±3.96) participated in this experiment.

Follow-up 'prediction' experiment (FU2): Sixty-nine individuals were recruited via Prolific Academic (www.prolific.ac; a crowdsourcing platform for research participation) to complete the experiment online and were paid at a fixed rate of £5.2 per hour. Two participants reported that they had taken part in a similar experiment before and were thus excluded prior to data analysis. The remaining 67 participants (42 males) were 27 years old (sd=±5.00) on average.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Materials

Food Images: All pictures were colour photographs selected from different sources on the Internet. They had a resolution of at least 512 X 384 pixels and were edited so that the single food item was presented in the centre against a white background using GNU Image Manipulation Program (GIMP) 2, a free open-source graphics editor. A selection of food was chosen based on the following criteria: first, the items would be familiar to participants to avoid hesitation due to uncertainty during decision-making; second, there was a wide variety of items including fruits/vegetables, sweets, snacks and savoury bites (e.g. grapes, salad, chocolate, nuts, sausage roll, etc.), which might accommodate different individual preferences and tastes for food and elicit different levels of desirability in participants.

Magic trick videos: Magic tricks, performed by three professional magicians including a champion of an international magic competition, were recorded in a TV studio with a professional cameraman using high resolution video cameras. All videos were then edited using Adobe® Premiere Pro CC® (2015) software to a similar monotonic (dark) background, size (720 x 404 pixels) and viewing focus. The videos were muted (and subtitles were added in a few videos, when needed). The face of the magician was hidden to avoid potential distraction due to appearance and any facial expressions (see Supplementary videos 1-2). Out of the pool of 166 video clips, we selected magic tricks to be used in the current study by ensuring that they (1) include a range of different features (such as the use of cards, sleight of hands, optical illusions) and (2) likely elicit curiosity to different extents (curiosity ratings were obtained in a different pilot study). The initial behavioural experiment included 45 food items and 45 magic trick videos. These videos ranged between 8 – 46 seconds long (mean=22.22; median=20). Due to the

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

precious scanning time allowed, the first fMRI experiment (fMRI experiment 1) used only a subset of 36 food items and 36 magic tricks from the initial experiment (length of the videos: range= 8-46 sec; mean=20.61; median=18.5). Then, the two follow-up experiments (FU1 & FU2) made use of the same set of stimuli as fMRI experiment 1.

Trivia questions: Sixty trivia questions were selected from a publicly-available 244-item database, created by our research group [28] (http://koumurayama.com/resources.php). The selected questions were not obviously culture or age specific. Also, for all questions, the answers were likely to be unknown to the majority of participants. The selection of items corresponded to different trivia categories that might elicit curiosity among individuals to different extents, including art/music, history/geography, movies/TV, nature/animals, science, space, sports, food, as well as other miscellaneous facts. To ensure within-person variability in curiosity evoked across the experiment, half of the chosen questions were picked among those with high mean curiosity scores in the database and the other half among those with low mean scores (rated by a sample of 1498 respondents from a separate online study; for more information, refer to [28]). The second fMRI experiment (fMRI experiment 2) used 60 trivia questions, as well as 60 food items. On average, the chosen questions contained 10 (ranging between 6-16) words.

Procedure

Curiosity-driven Decision-Making Task

The main task of the study followed similar procedures (although there were slight modifications for the two fMRI and the two follow-up experiments, which are detailed in Supplementary Methods). In brief, each trial started with a central fixation cross and then a brief letter cue ('M'

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

signified 'magic trick'; 'F' signified 'food') to prepare participants for the kind of stimulus they were about to see. This was followed by the stimulus (either a video of a magic trick in the curiosity condition or an image of a food item). Participants then gave a rating to indicate their level of curiosity about the magic trick or level of desirability of the food, on a 7-point scale (1=not at all, 7=very much). In the curiosity trials, they also had to report how confident they were to know about the solution to a magic trick, using also a 7-point scale (1=not at all, 7=very much). (However, because rated confidence was not associated with a participant's decision and the inclusion of this measure did not change any main results, to make straightforward comparisons between curiosity and food trials, it was not included in the reported analysis here.) After rating the stimulus, participants were then presented with a wheel of fortune (WoF) representing a lottery which visualised their probability of winning/losing in that trial, and were asked to decide whether to gamble or not. Participants were instructed that if they accepted the lottery and won, they would receive a token that might allow them to see the secret behind the magic trick after the experiment. If they gambled and lost, they would get a token that might increase the amount of shock they were to experience at the end of the experiment. Participants could also opt to reject the lottery. At the end of each trial, the outcome of the lottery was presented. Participants were informed that, as a general rule, the more 'win' tokens they collected, the more likely they would get more rewards. Similarly, the more 'loss' tokens they got, the more likely they would receive more electric shock.

There were 5 versions of WoF, each displaying a different combination of the probabilities of winning and losing a gamble: i) 16.7% (1/6) win vs. 83.3% (5/6) loss; ii) 33.3% (2/6) win vs. 66.7% (4/6) loss; iii) 50% (3/6) win vs. 50% (3/6) loss; iv) 66.7% (4/6) win vs. 33.3% (2/6) loss;

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

v) 83.3% (5/6) win vs. 16.7% (1/6) loss. In the experiments, participants were never shown the actual percentages but the relative win-to-loss contrast of probabilities was illustrated visually by the relative sizes of the constituent slices on the WoF (Fig. 1a). To control for the number of 'success' and 'loss' experiences, unbeknownst to the participants, there was an equal chance of winning or losing the lottery across all trials.

The curiosity and food trials were mixed and shown in a random order. The task was programmed and presented using PsychoPy [41].

Program on the day of the testing session

Participants were asked not to consume any food and drinks (apart from water) within at least 2 hours before attending their testing session. This was confirmed with the participants at the beginning of the study, including asking them to indicate when they last ate and had their last meal. Following standard procedures of informed consent and completing corresponding health and safety screening, participants underwent calibration for electric stimulation to identify a maximum (uncomfortable yet non-painful) threshold of electric shock they can endure (see Supplementary Methods for details). Then, they were told about the decision-making task. There was a practice task that used a different set of stimuli (3 curiosity, 3 food) prior to the main task.

At the end of the experiment, participants filled out a post-experiment questionnaire that asked the extent to which they expected to receive electric shock. In this questionnaire, the majority of participants across the two fMRI and the follow-up behavioural (i.e. FU1) experiments reported prospectively that during the experiment they expected to receive the electric shock - 89% of

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

them gave a rating of 3 or above out of 5 (mean=3.67, mode=4), representing their strong belief that the shock would have happened (we do not have this information from the initial behavioural study). After completing the questionnaire, rewards (including solutions to some magic tricks and food items) were delivered. Participants did not actually receive any electroshock. They were told that they were exempt from the shock as determined by a probabilistic equation programmed in the task.

fMRI acquisition

Whole-brain functional and anatomical images were acquired at the Centre for Integrative Neuroscience and Neurodynamics (CINN), University of Reading, UK using a 3.0 Tesla Siemens MAGNETOM Trio scanner with a 32-channel Head Matrix coil.

fMRI experiment 1 (Curiosity about Magic)

Functional images were acquired using a T2*-weighted gradient-echo echo planar imaging (EPI) pulse sequence with 37 axial slices (in-plane resolution of 3 x 3 x 3mm, interslice gap: 0.75mm), interleaved from bottom to top (echo time (TE): 30 ms; repetition time (TR): 2000 ms; flip angle: 90˚; field of view (FOV): 1344 x 1344 mm^2; in-plane matrix: 64 x 64). A high-resolution T1-weighted three-dimensional anatomical image was also collected, using an MPRAGE-gradient sequence with 176 x 1mm slices (in-plane resolution of 1 x 1 x 1 mm; TE: 2.52 ms; TR: 2020 ms; Inversion Time (TI):1100 ms; FOV: 250 x 250; flip angle: 9°), enabling optimal localisation of the functional effects.

fMRI experiment 2 (Curiosity about Trivia)

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Functional images were acquired using the same sequence and parameters as in fMRI experiment 1. A high-resolution T1-weighted three-dimensional anatomical image was collected using an MPRAGE-gradient sequence with 192 x 1 mm slices (TE: 2.29 ms; TR: 2300 ms; TI:900 ms; FOV: 240 x 240; flip angle: 8°).

fMRI analysis

Preprocessing and data analyses of the imaging data were performed using the SPM software (www.fil.ion.ucl.ac.uk/~spm). The preprocessing procedures included spatial realignment of the EPI volumes, co-registration with the structural image, segmentation, group-wise normalisation using DARTEL, and smoothing (see Supplementary Methods for more details).

Region-of-Interest (ROI) Mask

Given past findings and shown in our previous works [42,43], the striatum plays an important role in influencing motivation-driven behaviour. We created an anatomical mask encompassing the bilateral caudates and performed ROI analyses in fMRI experiment 1, using the WFU PickAtlas toolbox in SPM. The resulting anatomical mask included 832 voxels (2.5 resels).

To gauge the robustness and generalisability of the findings observed in fMRI experiment 1, we created a functional ROI mask based on the functionally activated voxels within the bilateral striatum from the ROI analysis of fMRI experiment 1 (specifically, from the main 'decision' effect: Acceptance > Rejection; thresholded at $p<0.001$, uncorrected; i.e., see fig 2b). This bilateral functionally defined caudate ROI, comprising 230 voxels (0.9 resels), was then used in the fMRI experiment 2 analysis. In particular, we computed the average beta/contrast values

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

within this functional mask using MarsBar [44] and conducted a single statistical test to examine whether the results of fMRI experiment 1 were replicated at a significance level of $P < 0.05$.

Activation predicting curiosity/incentive-driven decision-making

One of the aims of the current study was to test whether the participant's decision-making was influenced by curiosity (about knowledge) and food attraction in a similar manner via the incentive motivation system. To this aim, the first GLM modelled brain activation depending on the dichotomous decision in the lottery – whether an individual opted to accept or reject the gamble. Four separate regressors were specified for the accepted and rejected trials in curiosity and food categories at the onset of the decision phase of each trial (i.e. when the WoF appeared and the participant had to start indicating a choice). In addition, to account for the brain activation related to stimulus presentation, another 4 regressors were included to model the stimulus presentation phase (for magic trick trials specifically, this was time-locked at the moment of subjective surprise) of the accepted and rejected trials for the two stimulus categories (results not reported). We performed a 2x2 repeated-measures ANOVA with category (curiosity/food) and decision (accepted/rejected) as factors, at the decision phase. We hypothesised a main effect of decision (i.e. accepted > rejected) in the caudate ROI if curiosity and food influenced decision in a similar way via the incentive motivation system. We also examined the interaction effects to probe any potential stimulus-specific effects on decision-making. A familywise error-corrected voxel-level significance threshold of $P < 0.05$ was used within the ROI.

For additional whole brain analyses to explore activations outside of the defined ROI and in

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

other contrasts, we used a familywise error-corrected significance threshold of $P < 0.05$ at the cluster-level (cluster defining threshold: $P < 0.001$).

Controlling for the effect of the probability of outcome (winning/losing) probability

To factor out the effect of the presented probability of winning a lottery (or getting electric stimulation), we created another GLM and performed a parametric modulation analysis. In this GLM, the first-level design matrices included two main categorical regressors, modelling the onsets of the decision phase of trials for the two types of stimuli (curiosity, food). For these two regressors, a participant's decision in the lottery (i.e. 'accept' choice coded as +1; 'reject' choice coded as -1) was added as a parametric modulator (PM); the outcome probability was added as another PM. Similar to the other GLM, we still included reaction time as a PM to factor out the its potential confounding effect.

Conjunction analysis

To examine whether there were consistent effects in curiosity/incentive-driven decision-making regardless of the stimuli used, we performed a conjunction analysis on the results of the two fMRI studies (which made use of very different curiosity-inducing materials). The conjunction analyses, using a masking function of SPM, began with creating and saving a functional activation mask (image) resulting from a particular SPM contrast of interest (i.e. accepted > rejected trials) in the first fMRI experiment, thresholded at $P < 0.001$ (uncorrected). This activation image was then used as an inclusion mask for the same contrast in the second fMRI experiment (also thresholded at $P < 0.001$, uncorrected), resulting in a conjunction probability of $P < 0.00001$ for any 'activated' clusters within the mask.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Functional Connectivity Analysis

Functional connectivity was examined with the beta series correlation method [45] implemented in BASCO toolbox [46]. This method allows us to use trial-to-trial variability to characterise dynamic inter-regional interactions. Our a priori ROIs for functional connectivity analyses were the 6mm-radius spheres centred on the peak voxels of the left (MNI coordinate: -9, 15, 3) and right (6, 15, 3) caudate clusters identified in the main analysis (accepted > rejected gambles) of fMRI experiment 1.

At the first level of the analysis, a new GLM was constructed, in which BOLD response time-locked to the onset of the decision phase of each trial was modelled individually by a separate covariate using a canonical haemodynamic response function. This resulted in different parameter estimates for each trial for each participant. The onsets of stimulus presentation in each trial, and the six motion parameters for each run were also included in this GLM. Next, seed-based correlations were computed voxel-wise for each participant and for each of the experimental conditions of interest. This procedure generated an individual correlation map between each seed region's beta series and the beta series of all other voxels in the brain separately for each condition of interest, which was normalised using Fisher's r-to-z transformation. At the second level, the individual correlation maps were taken into random-effects t-tests to identify voxels that showed changes in functional connectivity with the ROI seed (based on trial-by-trial variability in parameter estimates) across different conditions. Given the exploratory nature of the analysis, in fMRI experiment 1, the analyses were thresholded at $P < 0.005$ at the voxel-level (uncorrected) (see Supplementary Table 4 for the results). In particular, a large cluster of voxels were revealed in the left sensorimotor cortex to be less

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

correlated (i.e. decoupled) with the caudate (spherical ROI) in accepted relative to rejected gambles (Fig. 3). We then created a functional ROI mask based on these 'functionally decoupled' sensorimotor cortex voxels (comprising 105 voxels) to be used in the replication analysis of fMRI experiment 2.

The functional connectivity analysis of fMRI experiment 2 mainly aimed to evaluate whether the results from fMRI experiment 1 could be replicated, using the same approach as the main GLM analysis. Specifically, with the correlation (functional connectivity) maps from fMRI experiment 2, we computed average contrast values (parameter estimates) within the functional sensorimotor cortex ROI using Marsbar toolbox, and performed a single statistical test to examine whether the functional connectivity results from fMRI experiment 1 (focusing on the differences in the average contrast values between 'accept' and 'reject' choices) could be replicated with the significance level set at $P < 0.05$. For further exploratory purposes, we performed an additional whole brain analysis of differences in functional connectivity between the 'accept' and 'reject' decisions in fMRI experiment 2 (see Supplementary Table 5).

Behavioural analysis

Linear mixed effect modelling

To examine the main question of whether curiosity and food desirability influence decisions on a trial-by-trial basis at the behavioural level, we performed analysis using a generalised linear mixed effects model (GLMM). In the model (separately for curiosity and food conditions), a participant's decision (i.e. choice in the lottery, which was either 'accept' or 'reject;) was specified as the dichotomous outcome, in a logistic link function. Rating of curiosity (or

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

desirability in the food condition) (group-mean centred) and the presented probability of winning/losing were entered as predictors of the decision. To account for the nested structure of the data, we specified random intercepts and slopes of participants.

We also combined the data from the initial behavioural experiment, fMRI experiment 1 and fMRI experiment 2 and applied the same GLMM to the combined dataset to synthesise the results (see Supplementary Table 1). We controlled for the effect of the experiment (i.e. from which experiment the data were taken from) as an extra fixed effect. This is equivalent to conducting a fixed-effect meta-analysis across the experiments. Importantly, the 'experiment' factor was not significant, indicating no differences in participants' performance (in terms of acceptance rate) between the three experiments.

The follow-up behavioural experiment FU1 had a specific aim to investigate whether curiosity and food desirability distort a participant's subjective probability of winning/losing. The main analysis in this experiment was based on a linear mixed effects model (LMM) (again, separately for curiosity and food conditions) that included the rating of curiosity (or desirability in the food condition) and the presented probability of winning/losing as predictors of subjective probability of winning the lottery. Again, we specified random intercepts and slopes of participants. We compared two models, one including rated curiosity (or desirability in the food condition) and one without, and computed Bayes factors based on the Bayesian Information Criterion from each model to evaluate to the effect of the rating. Conventional cut-offs for interpreting Bayes factors are typically based on the ones suggested by Harold Jeffreys[47].

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The follow-up 'prediction' experiment FU2 had participants to predict what choices they thought other people would make. To compare and test whether these predictions were different from (i.e. underestimating) the actual decisions made by the 'real-life' participants, we performed a GLMM incorporating data from the initial behavioural experiment, the two fMRI experiments, as well as the "prediction" study FU2. In this model, there was a fixed-effect factor 'prediction' which specified whether a response was a prediction (+1) or an actual decision (-1), a factor 'rating' which specified participants' ratings of curiosity/desirability of food, a 2-way interaction of these factors, as well as a factor 'category' and a factor for 'the presented probability of lottery outcome' to control for these effects. Importantly, the prediction X rating interaction effect was negative and significant – specifically, the effect of rating was smaller in the 'prediction' experiment than in the other actual experiments, suggesting that making a prediction for others underestimates the actual effect of curiosity and food desirability in overcoming fear of physical risks (see Supplementary Table 6).

Testing of GLMM and LMM was carried out using the package 'lme4' in R.

Mediation analysis

To test whether the relationship between rated curiosity/food desirability and the decision to accept a lottery is mediated by the striatal activation observed in the main fMRI analysis, we performed a multilevel-mediation analysis using multilevel structural equation modelling with Mplus [48] (version 7).

In this analysis, we specified the indirect paths as starting from Rating through Striatal_activation

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

to Decision (i.e. Rating → Striatal activation → Decision) as well as the direct path from Rating to Decision, and analysed data of curiosity and food conditions altogether using trials as the unit of analysis. Category of the stimulus (i.e. curiosity or food) and the presented probability of winning were included as covariates in the model. Trial-by-trial average striatal activation was extracted within the ROI masks, 6mm-radius spheres centred on the peak voxels of the left (MNI coordinates: -9, 15, 3) and right (6, 15, 3) caudate clusters (i.e., the cluster of the largest activation) identified previously in the main analysis (accepted > rejected gambles) in fMRI experiment 1. Decision was treated as a binary outcome variable.

To account for the multilevel nature of the data, we used person-mean centring [49] before parameter estimation and computed cluster robust-standard error to make statistical inference [50]. After estimating the model, mediation effect was tested by directly evaluating the distribution of the product confidence limits for the mediation effect using the software PRODCLIN [51]. We claimed statistical significance at 5% if the 95% confidence interval computed by PRODCLIN did not include zero.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

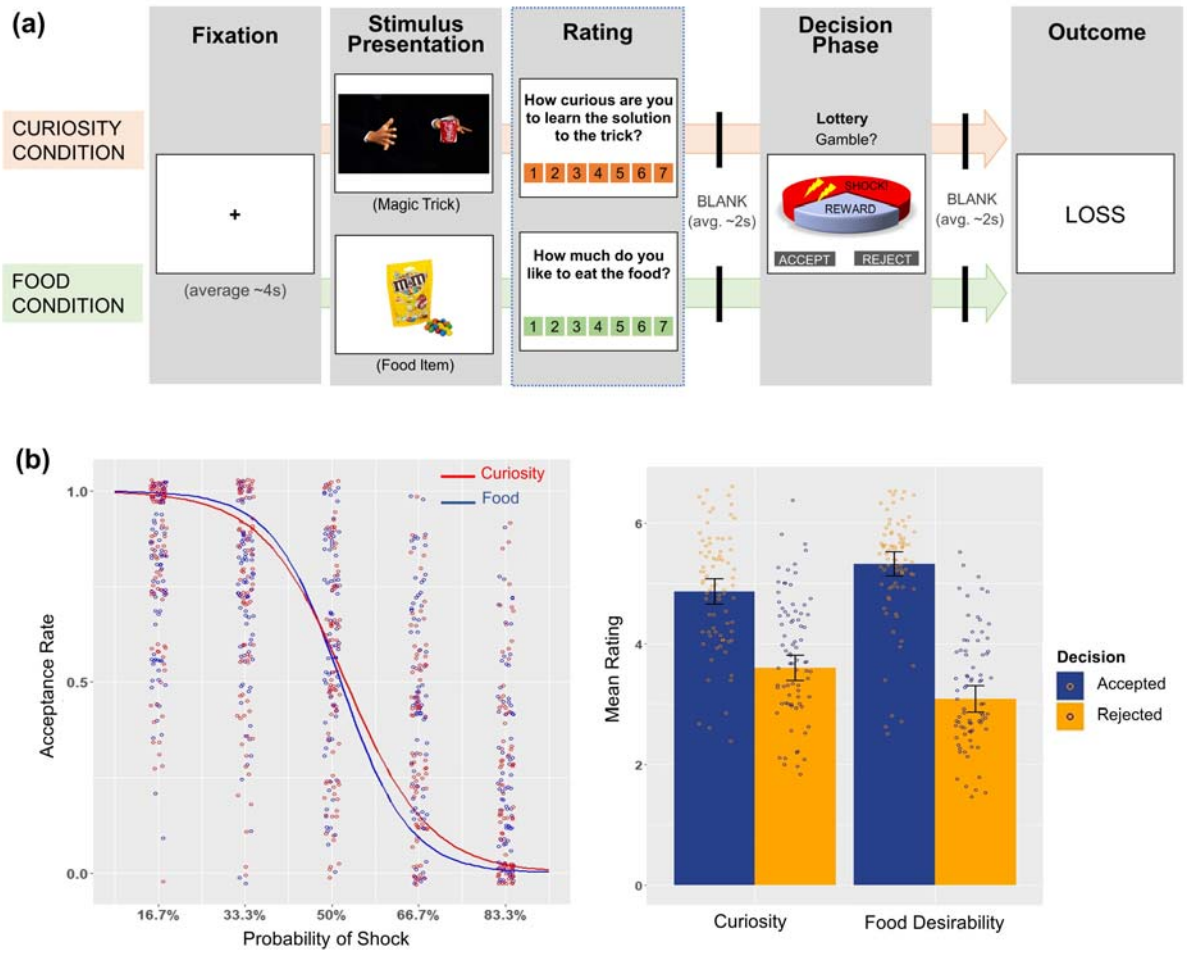

Fig. 1 | The experimental task and behavioural results of motivation-driven decision-making. a, Illustration of a typical trial of the experimental task in curiosity and food conditions. In modified versions for the fMRI experiments, (due to time constraints in scanning) the rating phase appeared in only 10% of all trials during the task. **b,** Behavioural results integrated from the initial behavioural and two fMRI experiments. *On the left:* Modelled logistic curve illustrates participants' decisions to accept the lottery at different levels of probability of electric shock in food (blue curve) and curiosity (red curve) conditions. Acceptance rate on y-axis ranges from 0 (100% rejection) to 1 (100% acceptance). Each dot represents the average acceptance rate of a participant at that level of shock probability. *On the right:* Bar chart shows higher average rating

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

of curiosity (about magic tricks and trivia questions) and food desirability in accepted (blue) compared with rejected (orange) gambles. Each dot represents the average rating per participant, for each condition. Error bars represent ±1 standard error of the mean (SEM). Dots on the graphs are jittered for better visual display. Also see Supplementary Fig. 1 for illustration of the same effects in individual studies separately.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

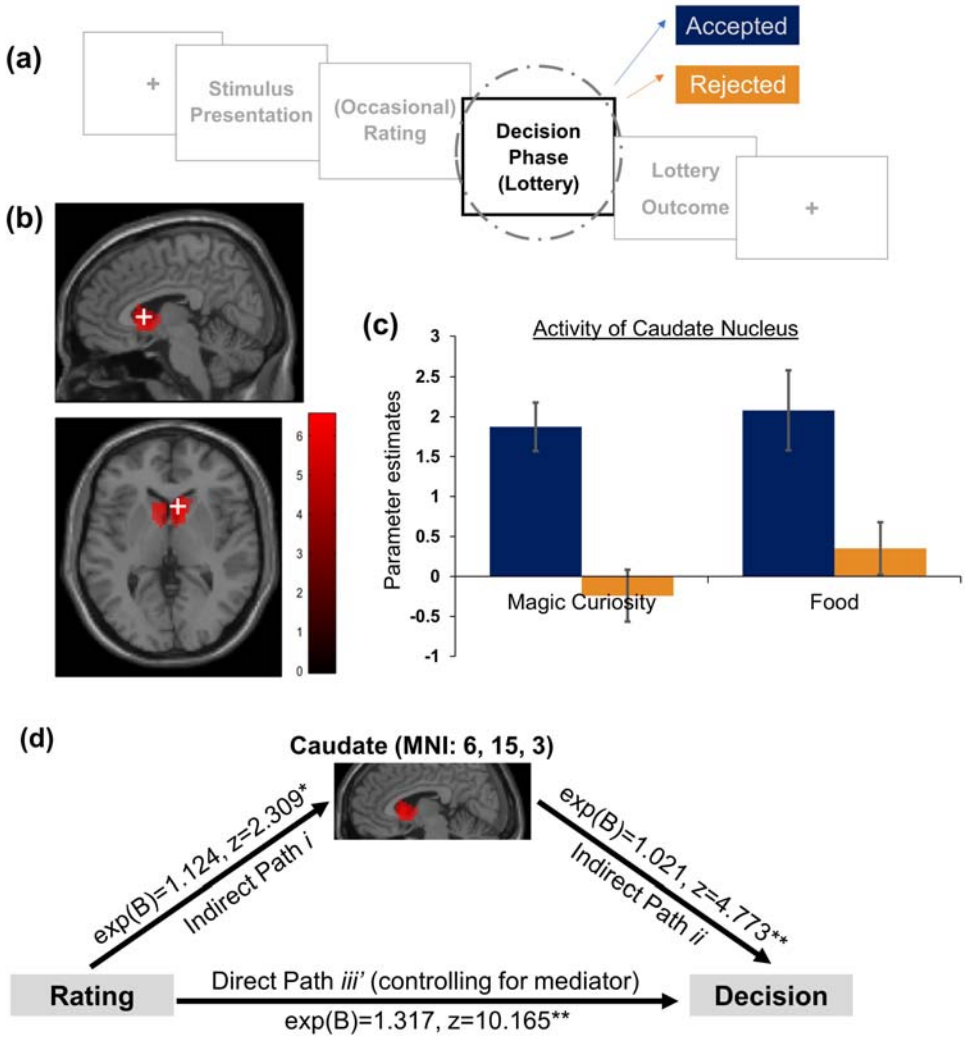

Fig. 2 | Motivation-driven decision-making is mediated by the striatum (fMRI experiment 1). a, Analysis of brain activity associated with decision-making was based on whether a participant accepted or rejected the lottery. **b,** Increased activity within the bilateral caudate nucleus for accepted compared with rejected gambles at the decision phase was significant at $P<0.05$ with voxel-wise FWE-correction. For display, the results shown here (in red) are thresholded at $P<0.001$ (uncorrected). **c,** The bar plot shows the mean parameter estimates (i.e. brain activations) for a 6mm-radius sphere centred at a peak voxel (MNI: 6, 15, 3) of the caudate cluster (indicated by the white crosses in **b**) across participants. Error bars represent ±1 SEM for each condition. **d,** In the mediation path diagram, rating (of magic curiosity and food

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

desirability) predicts the mean activation of the sphere in the caudate nucleus, illustrated by Path *i*. The connection of caudate activation (the mediator) to decision in the lottery (the outcome), Path *ii*, was calculated controlling for rating, as is standard practice for mediation models. The direct path is Path *iii'*, calculated controlling for the mediator. The lines are labelled with exponential values of path coefficients, exp(B), and z values. This model also controlled for the effects of the category of the stimulus and the presented outcome probability in each trial. *significant at $P<0.05$; **$P<0.01$.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

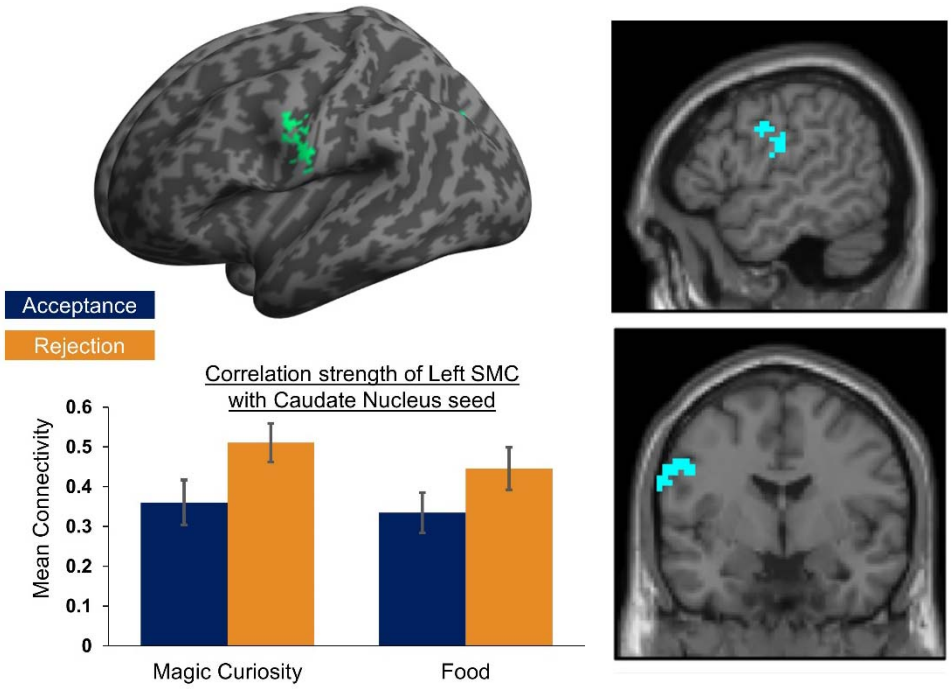

Fig. 3 | Functional connectivity of caudate nucleus in fMRI experiment 1. Analysis of whole-brain functional connectivity was based on the beta series correlation method using the trial-by-trial variability[45]; the ROI was defined as a 6mm-radius sphere centred at the peak voxel (MNI: -9, 15, 3) in the caudate nucleus (left) identified in the main fMRI analysis. This connectivity analysis revealed a significant main effect of decision in the left sensorimotor cortex (SMC), in which activation was less correlated with the caudate ROI in the accepted compared with rejected gambles (thresholded at $P<0.005$, uncorrected). This is illustrated in 3D rendered and whole brain images in the figure. The bar plot shows the magnitude of correlation (r, converted back from z values) between activities in the left caudate ROI and the left SMC cluster in different trial types. Error bars indicate ±1 SEM.

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

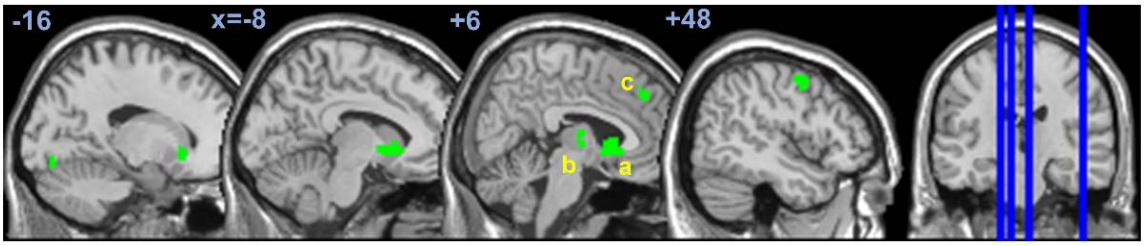

Fig. 4 |Common effects of motivation-driven decision from the two fMRI experiments. The sagittal multi-slice brain images show the conjunction of activations from the main GLM whole-brain analysis of fMRI experiment 1 & 2, each thresholded at the statistical significance of P<0.001 (uncorrected), giving rise to a conjunction probability of *P*<0.00001. Shown in green, the conjunction effects for the contrast 'Accepted > Rejected gambles' (across both stimulus types) are found in the left and right caudate nuclei (in slices x=-8 and x=6 (cluster a) respectively), right thalamus (cluster b), dorsal medial frontal lobe (cluster c), and right premotor cortex (x=48).

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

References

1. von Stumm, S., Hell, B. & Chamorro-Premuzic, T. The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspect. Psychol. Sci.* **6,** 574–588 (2011).
2. Gruber, M. J., Gelman, B. D. & Ranganath, C. States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* **84,** 486–496 (2014).
3. Kang, M. J. *et al.* The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychol. Sci.* **20,** 963–973 (2009).
4. Gottlieb, J., Oudeyer, P.-Y., Lopes, M. & Baranes, A. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends Cogn. Sci.* **17,** 585–593 (2013).
5. Jirout, J. & Klahr, D. Children's scientific curiosity: In search of an operational definition of an elusive concept. *Dev. Rev.* **32,** 125–160 (2012).
6. Kidd, C. & Hayden, B. Y. The psychology and neuroscience of curiosity. *Neuron* **88,** 449–460 (2015).
7. Renninger, K. A. & Hidi, S. *The power of interest for motivation and engagement*. (Routledge, NY, 2016).
8. Sakaki, M., Yagi, A. & Murayama, K. Curiosity in old age: A possible key to achieving adaptive aging. *Neurosci. Biobehav. Rev.* **88,** 106–116 (2018).
9. Rozek, C. S., Svoboda, R. C., Harackiewicz, J. M., Hulleman, C. S. & Hyde, J. S. Utility-value intervention with parents increases students' STEM preparation and career pursuit. *Proc. Natl. Acad. Sci.* **114,** 909–914 (2017).
10. Loewenstein, G. The psychology of curiosity: A review and reinterpretation. *Psychol. Bull.* **116,** 75–98 (1994).

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

11. 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).

12. 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).

13. Vasconcelos, M., Monteiro, T. & Kacelnik, A. Irrational choice and the value of information. *Sci. Rep.* **5,** 13874 (2015).

14. Bennett, D., Bode, S., Brydevall, M., Warren, H. & Murawski, C. Intrinsic valuation of information in decision making under uncertainty. *PLOS Comput. Biol.* **12,** e1005020 (2016).

15. Brydevall, M., Bennett, D., Murawski, C. & Bode, S. The neural encoding of information prediction errors during non-instrumental information seeking. *Sci. Rep.* **8,** 6134 (2018).

16. Eliaz, K. & Schotter, A. Paying for confidence: An experimental study of the demand for non-instrumental information. *Games Econ. Behav.* **70,** 304–324 (2010).

17. Wright, W. F. & Bower, G. H. Mood effects on subjective probability assessment. *Organ. Behav. Hum. Decis. Process.* **52,** 276–291 (1992).

18. Litman, J. Curiosity and the pleasures of learning: Wanting and liking new information. *Cogn. Emot.* **19,** 793–814 (2005).

19. Beckmann, J. S. & Chow, J. J. Isolating the incentive salience of reward-associated stimuli: value, choice, and persistence. *Learn. Mem.* **22,** 116–127 (2015).

20. Berridge, K. C. From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* **35,** 1124–1143 (2012).

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

21. Berridge, K. C. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl).* **191,** 391–431 (2007).

22. Berridge, K. C., Robinson, T. E. & Aldridge, J. W. Dissecting components of reward: 'liking', 'wanting', and learning. *Curr. Opin. Pharmacol.* **9,** 65–73 (2009).

23. Kringelbach, M. L. & Berridge, K. C. in *Recent Developments in Neuroscience Research on Human Motivation* (eds. Kim, S., Reeve, J. & Bong, M.) 23–35 (Emerald Group Publishing Limited, Bingley, 2016). doi:10.1108/S0749-742320160000019020

24. Delgado, M. R., Schotter, A., Ozbay, E. Y. & Phelps, E. A. Understanding overbidding: Using the neural circuitry of reward to design economic auctions. *Science*. **321,** 1849–1852 (2008).

25. Hayes, D. J. *et al.* GABAA receptors predict aversion-related brain responses: An fMRI-PET investigation in healthy humans. *Neuropsychopharmacology* **38,** 1438–1450 (2013).

26. Whitmarsh, S., Nieuwenhuis, I. L. C., Barendregt, H. P. & Jensen, O. Sensorimotor alpha cctivity is modulated in response to the observation of pain in others. *Front. Hum. Neurosci.* **5,** (2011).

27. Berlyne, D. E. A theory of human curiosity. *Br. J. Psychol.* **45,** 180–191 (1954).

28. Fastrich, G. M., Kerr, T., Castel, A. D. & Murayama, K. The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motiv. Sci.* **4,** 227–250 (2018).

29. Loewenstein, G. & Lerner, J. S. in *Handbook of Affective Science* (eds. Davidson, R., Goldsmith, H. & Scherer, K.) 619–642 (Oxford University Press, Oxford, 2003).

30. Gilbert, D. T., Lieberman, M. D., Morewedge, C. K. & Wilson, T. D. The peculiar longevity of things not so bad. *Psychol. Sci.* **15,** 14–19 (2004).

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

31. Murayama, K., FitzGibbon, L. & Sakaki, M. Process account of curiosity and interest: A reward learning model of knowledge acquisition (OSF Preprints). Preprint at https://osf.io/hbcz5 (2018).

32. Hsee, C. K. & Ruan, B. The Pandora effect: The power and peril of curiosity. *Psychol. Sci.* **27,** 659–666 (2016).

33. Wilson, T. D. *et al.* Just think: The challenges of the disengaged mind. *Science.* **345,** 75–77 (2014).

34. Kruger, J. & Evans, M. The paradox of Alypius and the pursuit of unwanted information. *J. Exp. Soc. Psychol.* **45,** 1173–1179 (2009).

35. Oosterwijk, S. Choosing the negative: A behavioral demonstration of morbid curiosity. *PLoS One* **12,** e0178399 (2017).

36. Mobbs, D. *et al.* When fear Is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science.* **317,** 1079–1083 (2007).

37. Adolphs, R. The biology of fear. *Curr. Biol.* **23,** R79-93 (2013).

38. Dubois, J., Galdi, P., Han, Y., Paul, L. K. & Adolphs, R. Resting-state functional brain connectivity best predicts the personality dimension of openness to experience. *Personal. Neurosci.* **1,** e6 (2018).

39. Oudeyer, P.-Y. & Kaplan, F. What is intrinsic motivation? A typology of computational approaches. *Front. Neurorobot.* **1,** 1–14 (2007).

40. Pisula, W., Turlejski, K. & Charles, E. P. Comparative psychology as unified psychology: The case of curiosity and other novelty-related behavior. *Rev. Gen. Psychol.* **17,** 224–229 (2013).

41. Peirce, J. W. Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.* **2,**

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(2009).

42. Murayama, K., Matsumoto, M., Izuma, K. & Matsumoto, K. Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proc. Natl. Acad. Sci.* **107,** 20911–20916 (2010).

43. Murayama, K. *et al.* How self-determined choice facilitates performance: A Key role of the ventromedial prefrontal cortex. *Cereb. Cortex* **25,** 1241–1251 (2015).

44. Brett, M., Anton, J.-L., Valabregue, R. & Poline, J.-B. Region of interest analysis using an SPM toolbox. In: 8th International Conference on Functional Mapping of the Human Brain. *NeruoImage* **16,** (2002).

45. Rissman, J., Gazzaley, A. & D'Esposito, M. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* **23,** 752–763 (2004).

46. Göttlich, M., Beyer, F. & Krämer, U. M. BASCO: a toolbox for task-related functional connectivity. *Front. Syst. Neurosci.* **9,** (2015).

47. Jeffreys, H. *Theory of Probability*. (Clarendon Press, Oxford, 1961).

48. Muthén, L. K. & Muthén, B. O. *Mplus User's Guide. Seventh Edition.* (2012).

49. Enders, C. K. & Tofighi, D. Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychol. Methods* **12,** 121–138 (2007).

50. McNeish, D., Stapleton, L. M. & Silverman, R. D. On the unnecessary ubiquity of hierarchical linear modeling. *Psychol. Methods* **22,** 114–140 (2017).

51. MacKinnon, D. P., Fritz, M. S., Williams, J. & Lockwood, C. M. Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behav. Res. Methods* **39,** 384–389 (2007).

bioRxiv preprint doi: https://doi.org/10.1101/473975; this version posted November 22, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Acknowledgments

The study was supported by the Marie Curie Career Integration Grant (CIG630680 to K.M.), JSPS KAKENHI (15H05401; 16H06406 , 18H01102, and 18K18696 to K.M.), F. J. McGuigan Early Career Investigator Prize (to K.M.), and the Leverhulme Trust (RPG-2016-146 and RL-2016-030 to K.M.). We are grateful of the magicians (including Shota Irieda and Ohkubo Kohei) for producing magic tricks for our research. We thank E. Daveau for helping with data collection, C. Inaltay, A. Firat, A. Haffey, J. Raw, and G. Fastrich for editing and pilot-testing magic video clips, and C. McNabb for providing useful comments on the drafts of the article.

Author contributions

K.M. conceived the idea; J.K.L.L. and K.M. jointly designed the study; H.O., A.K., and K.K. created experimental materials. J.K.L.L. performed research and analyzed the data; J.K.L.L. and K.M. jointly wrote the paper. All authors provided critical comments. The authors declare no conflicts of interest.

Supplementary Materials:

Supplementary Methods

Supplementary Figures 1-4

Supplementary Tables 1-6

Captions for Supplementary videos