

# Anterior insula reflects surprise in value-based decision-making and perception



Leyla Loued-Khenissi<sup>a,\*</sup>, Adrien Pfeuffer<sup>b</sup>, Wolfgang Einhäuser<sup>c</sup>, Kerstin Preuschoff<sup>d,e</sup>

<sup>a</sup> Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, 1015, Switzerland

<sup>b</sup> Department of Physics, University of Marburg, Marburg, 35037, Germany

<sup>c</sup> Institute of Physics, Chemnitz University of Technology, Chemnitz, 09111, Germany

<sup>d</sup> Swiss Center for Affective Sciences, University of Geneva, Geneva, 1202, Switzerland

<sup>e</sup> Geneva Finance Research Institute, University of Geneva, Geneva, 1211, Switzerland

## ARTICLE INFO

### Keywords:

Insula  
Decision-making  
Uncertainty  
Inference  
fMRI  
Necker cube

## ABSTRACT

The brain has been theorized to employ inferential processes to overcome the problem of uncertainty. Inference is thought to underlie neural processes, including in disparate domains such as value-based decision-making and perception. Value-based decision-making commonly involves deliberation, a time-consuming process that requires conscious consideration of decision variables. Perception, by contrast, is thought to be automatic and effortless. Both processes may call on a general neural system to resolve for uncertainty however. We addressed this question by directly comparing uncertainty signals in visual perception and an economic task using fMRI. We presented the same individuals with different versions of a bi-stable figure (Necker's cube) and with a gambling task during fMRI acquisition. We experimentally varied uncertainty, either on perceptual state or financial outcome. We found that inferential errors indexed by a formal account of surprise in the gambling task yielded BOLD responses in the anterior insula, in line with earlier findings. Moreover, we found perceptual uncertainty and surprise in the Necker Cube task yielded similar responses in the anterior insula. These results suggest that uncertainty, irrespective of domain, correlates to a common brain region, the anterior insula. These findings provide empirical evidence that the brain interacts with its environment through inferential processes.

## 1. Introduction

Uncertainty is an unavoidable feature of decision-making. This suggests the brain acts as an inference machine, formulating hypotheses about its environment and testing them against the evidence (Gregory, 1997; Clark, 2013). The idea that inference may underlie perception, an unconscious, rapid decision, was posited by the 11th century scientist Alhazen (Howard, 1996) and most famously stated by Hermann Von Helmholtz (1867) in the 19th century, who coined the term unconscious inference to describe the process. By contrast, in value-based decision-making, where options are quantified and mulled over by at least one conscious agent, the inferential process includes an explicit, objective component; agents commonly declare their uncertainty in quantitative terms, such as the odds of winning a bet. While several factors differentiate perception from value-based decision-making, they can both be cast as inferential processes, a computation that integrates and resolves for uncertainty. The key question arises: do similar neural mechanisms

process uncertainty in value-based decisions and perception, irrespective of decision features or goals (Parr and Friston, 2017)?

In casting the brain as an inference machine (Friston and Stefan, 2007; Doya et al., 2007), agents are theorized to build internal models that forecast outcomes (Friston, 2010; Clark, 2013). When the model's prediction fails, an error signal is generated that refines subsequent expectations. Model building and updating is not limited to deliberative judgments, but extends to latent decision-making, as in perception (Haefner et al., 2016). A shift in perception of an unchanging stimulus for instance hints at an internal inferential error (Von Helmholtz, 1867). If uncertainty drives an agent to infer an outcome, predictions on uncertainty are also made. This prediction uncertainty is dubbed expected uncertainty (Payzan-Le Nestour et al., 2013), or risk (Preuschoff et al., 2006; Faraji et al., 2018) and represents the "known unknowns". Errors in uncertainty prediction are dubbed surprise (Preuschoff et al., 2011). Both risk and surprise can be quantified by formal accounts. As predictions and prediction errors arise from uncertainty, decisions cast in a

\* Corresponding author. Theory of Pain Laboratory, University of Geneva, Geneva, Switzerland.

E-mail address: [lkhenissi@gmail.com](mailto:lkhenissi@gmail.com) (L. Loued-Khenissi).

<https://doi.org/10.1016/j.neuroimage.2020.116549>

Received 21 May 2019; Received in revised form 20 December 2019; Accepted 14 January 2020

Available online 16 January 2020

1053-8119/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

probabilistic framework formally encapsulate an inferential process. Measuring uncertainty signals in the brain presents one way to index inference, irrespective of temporal dynamics, consciousness features, or end-goal of a decision.

Perception rarely requires deliberation. Nonetheless, uncertainty arises, as the mapping of the distal world to the sensory signal is not uniquely invertible. Perception infers the world by combining sensory evidence with prior knowledge (Von Helmholtz, 1867), an approach often cast in Bayesian frameworks (Freeman, 1994; Knill and Pouget, 2004). Perceptual uncertainty has traditionally been captured by drift-diffusion (Ratcliff, 1979) and signal detection models (Heekeren et al., 2008), where reaction time distributions serve to identify latent variables including decision difficulty, or stimulus uncertainty. In such paradigms, a volitional decision, usually motor, is made when uncertainty is resolved. With multi-stable stimuli however, information conveyed by a constant stimulus is sufficiently uncertain to evoke spontaneous alternations in conscious perception (Rubin, 1921; Boring, 1930; Necker, 1832), suggesting the decision process itself is hidden, with only the outcome made available to the agent. Crucially, ambiguous perception does not rely on dominance time as an index of uncertainty (Brascamp et al., 2005), as switches arise in a stochastic manner. Bi-stable figures have recently been studied in predictive-coding frameworks where perceptual transitions (see also: Dayan, 1998; Hohwy et al., 2008) correlated with anterior insulae and inferior frontal gyri responses (Weilhammer et al., 2017), regions also associated with economic uncertainty (Platt and Huettel, 2008; Grinband et al., 2006; Brevers et al., 2015; Preuschoff et al., 2008a; Mohr et al., 2010).

Studies on perception and value-based decision-making have generally remained distinct (Summerfield and Tsetsos, 2012) though some have sought modality-independent neural correlates of the decision process (Ho et al., 2009; Heekeren et al., 2006). Value-based decision-making is driven by utility maximization (Von Neumann and Morgenstern, 1944), while successful perception aims for an accurate mental representation of a stimulus. Expected utility violations entail a reward prediction error (Schultz et al., 1997); unexpected shifts in perception should entail perceptual prediction errors (Egner et al., 2010). These domain-specific errors may correlate with distinct neural systems. However, neither utility nor perceptual prediction can elude uncertainty, a quantity whose minimization is posited to drive decision-making (Friston et al., 2017; Schwartenbeck et al., 2015). Thus percept variance is analogous to the risk in predicting the utility of an investment (Markowitz and March, 1952). Similarly, as errors in predicting uncertainty arise in financial paradigms (Preuschoff et al., 2006, 2008a), so they should in perception, when uncertainty breaks through the consciousness barrier, as with bi-stable stimuli. This error in uncertainty prediction can be dubbed surprise (Preuschoff, 2006; Schwartenbeck et al., 2015; Faraji et al., 2018).

In previous work, we studied objective, expected uncertainty (risk) and its error (surprise) in value-based decision-making with a gambling task, finding a role for the insula in response to both (Preuschoff et al., 2008a). To determine if the same computational account and neural mechanism capture perceptual uncertainty and surprise, we sought a task that mimics the probabilistic nature of a gamble, but provokes internally generated perceptual errors. To meet this last criterion, we elicited uncertainty with a multi-stable stimulus, the Necker Cube (Necker, 1832), which prompts spontaneous perceptual switches without a corresponding change in stimulus.

The insula is a bilateral, cortical region that is generally involved in interoception, where bodily signals are integrated into feeling states (Craig and Craig, 2009). The region is implicated in several functions (Uddin et al., 2017) but is commonly found in studies that include error detection (Klein et al., 2013) or a deviant stimulus (Menon and Uddin, 2010). This hypothesized function of the anterior insula highlights its eligibility as a candidate region for generalized uncertainty processing, as uncertainty presents both a mathematical error and a feeling state. Given the anterior insula's emergence in perceptual difficulty (Binder et al.,

2004; Thielscher and Pessoa, 2007) and in economic uncertainty (Platt and Huettel, 2008; Grinband et al., 2006; Brevers et al., 2015; Preuschoff et al., 2008a; Mohr et al., 2010; Rutledge et al., 2010), we hypothesized this region's responses correlate with domain-independent uncertainty (Preuschoff et al., 2008b).

Does uncertainty call on a dedicated neural system, irrespective of functional domain? We examine this question with a common computational account of uncertainty in both perceptual and financial tasks, administered to the same participants during functional magnetic resonance (fMRI) acquisition. We hypothesized that 1) first order perceptual and reward prediction errors prompt distinct, task-specific BOLD responses in visual and striatal areas, respectively; 2) endogenous (perceptual) and exogenous (economic) uncertainty, will correlate with an insular response.

## 2. Materials and methods

### 2.1. Recruitment

The study was performed across two sessions scheduled within one week of each other. Sessions were separated due to the considerable length of time spent in the scanner for each (~45 min). The first session comprised the financial uncertainty task (the Card Game, Fig. 1a) and the second the perceptual uncertainty task (the Necker Cube Onset task and the Necker Cube Continuous Task) (Fig. 1c and d). A total of 29 participants were recruited in the Card Game (13 female, 16 male) with a mean age of 25.13 years, of which 4 participants' datasets were excluded for the following reasons; behavioral datasets from the three first participants were not recorded due to technical reasons, while the 4th showed an error rate (combination of missed bets and incorrect trial outcomes) greater than 30% across the three sessions. Twenty-two of the initial pool of 29 participants went on to perform the Necker Cube task; the other 7 were lost to attrition. Of these, 19 had also completed the Card Game. In the Necker Cube onset task, 22 subjects were included for analysis (one being excluded for poor behavioral responses to stimulus classifications) and in the Necker Cube continuous task, one additional subject was excluded for inconsistent behavioral data (too few perceptual switches across the experiment). In total, 18 same participants (8 female, 10 male) completed both sessions to yield useable neuroimaging and behavioral data. The study was approved by the local ethics committee and conformed to the declaration of Helsinki. Participants were recruited through online and paper advertisements broadcast on the Ecole Polytechnique Fédérale de Lausanne and Université de Lausanne campuses. The following inclusion criteria were applied: English speaking; over the age of 18; healthy; normal vision. Exclusion criteria from participation included: history of psychiatric or neurological illness; previous or current psychotropic drug use; metal implants; pregnancy; sensitivity to noise or closed spaces.

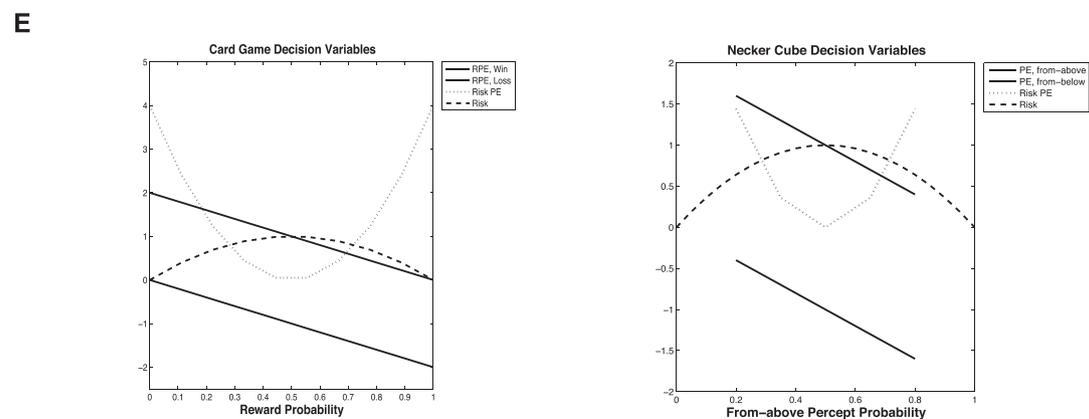
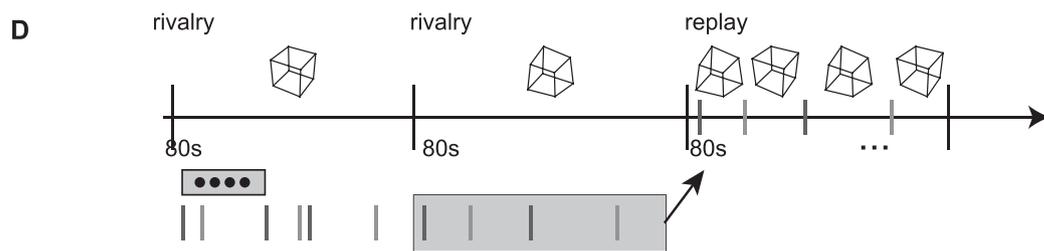
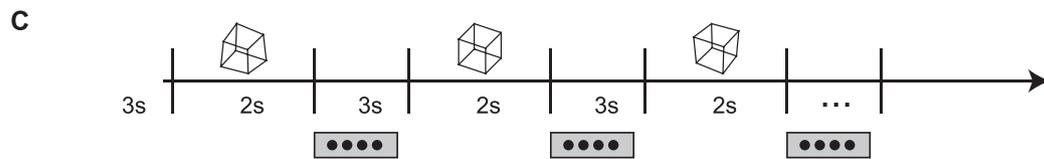
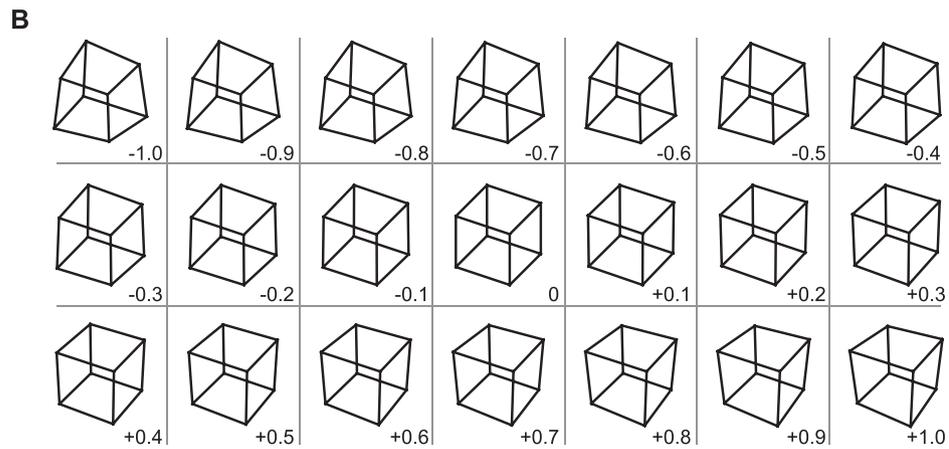
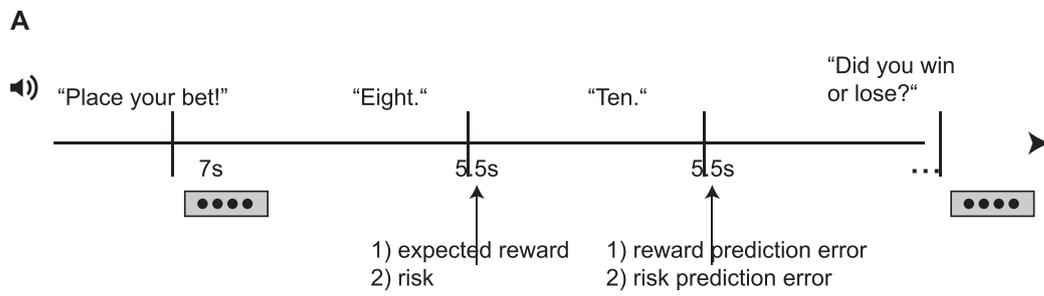
### 2.2. Procedure

For each of the two sessions, participants were sent digital versions of the study's information and consent forms on the eve of their study sessions for review. Upon arrival to the study room, participants were queried regarding their understanding of the information before providing their signed consent. Participants were then subject to additional MR safety screening prior to entering the scanning room.

### 2.3. Tasks

#### 2.3.1. Financial uncertainty task

We employed a gambling task performed during functional magnetic resonance imaging (fMRI) acquisition. We used an auditory version of a card game (Preuschoff et al., 2006), which has yielded reliable results when modeled by our chosen computational framework. In the task, participants had an equal probability of winning or losing 1 CHF (~1



(caption on next page)

USD) at each trial, starting with an initial endowment of 25 CHF. At each round, participants were instructed to place a bet via manual button press on whether a second card drawn from a deck of ten cards would be higher or lower than a first card drawn from the same deck. The bet was made prior to any card being drawn, at which point outcome probability is at 0.5, and the expected value of reward is 0. Following the bet, participants heard the value of the first card. After a 5.5 s interval, participants heard the value of the second card. After another 5.5 s interval, participants were instructed to report whether they had won or lost the round. An incorrect response incurred a penalty of 25 ¢ off a round's total payoff. Each trial lasted approximately 25 s (Fig. 1a). Inter-trial interval durations were randomly jittered (2–5s). All possible card pairs, excluding pairs of identically valued cards, were presented to each participant in a random order, totaling 90 trials per experiment. No two participants played the same sequence of gambles. These 90 trials were divided into three blocks of 30 trials each, to give participants the opportunity to rest between blocks. Each block began with a new 25 CHF endowment. As the task was auditory in nature, once the functional sequence was launched, participants were presented with a black fixation-cross centered on a gray-scale screen throughout the experiment. The task was sounded with the use of Mac OSX text-to-speech function, with the voice of 'Alex'. Instructions and stimuli were pre-recorded into wav files and transmitted to the participant via MR compatible headphones. The task related session lasted approximately 36 min while the imaging session lasted approximately 50 min; the whole experimental session, including intake and debriefing lasted approximately 90 min.

### 2.3.2. Perceptual uncertainty tasks

To manipulate perceptual uncertainty, we used different versions of a "Necker Cube" stimulus (Fig. 1b). Most individuals perceive this wire-frame representation of a cube as either a cube seen from-above or from-below, with alternations between the two interpretations. We warped the angles of the cube to bias perceptual interpretations towards one or the other direction.

**2.3.2.1. Necker Cube, onset task.** We presented observers with the 21 versions of the cube depicted in Fig. 1b. Each version was presented a total of 10 times, in random order (Fig. 1c). Each presentation lasted 3s, followed by a 2-s blank. Observers were asked to report their percept ("from-above", "from-below", as instructed with pictures of extreme versions prior to the experiment) by using two buttons of an MR-compatible response box. To avoid lateralization effects, the assignment of buttons to percepts was counterbalanced across participants. If participants failed to respond within the 2s blank period, the trial was treated as missing data.

**2.3.2.2. Necker Cube, continuous task.** For this task, 5 different versions of the Necker cube were used. These were selected for each individual based on onset-task results. To this end, we fitted a cumulative Gaussian to the dependence of the fraction of "seen from-above" reports on the physical distortion level (cf. Fig. 2a). Based on this fit, we determined the

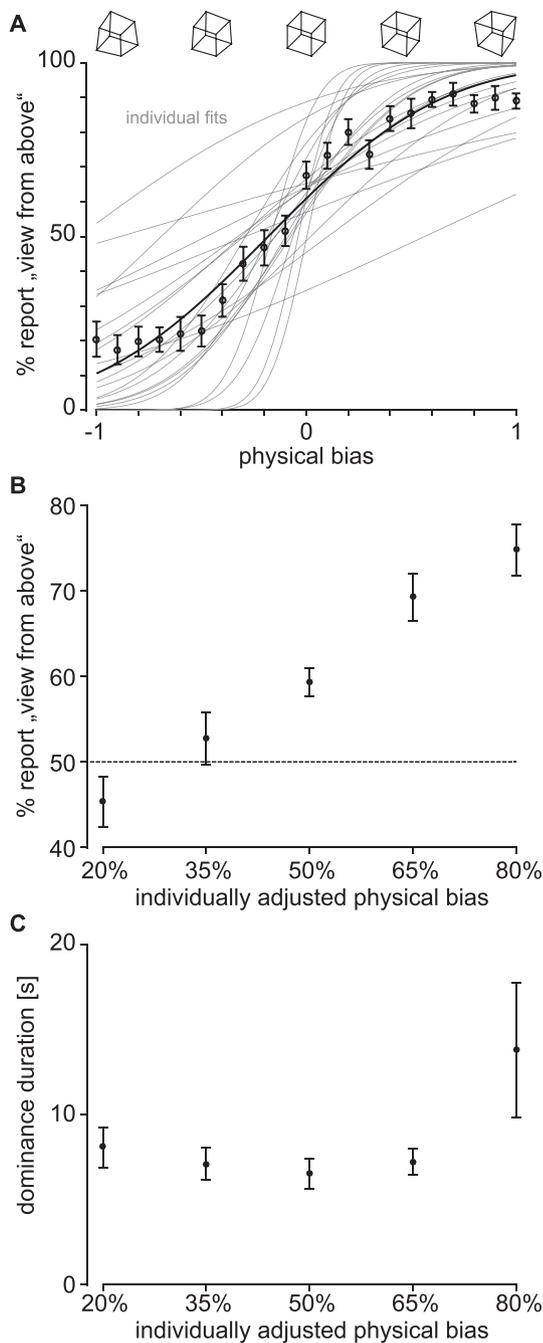
points at which 20%, 35%, 50%, 65% and 80% views from-above would be reported. (Note that in a few cases, this implied extrapolation to more extreme warpings). Although it cannot be expected that the predominance of the "from-above" view during prolonged viewing (i.e., the percentage of time the "from-above" view is reported dominant) provides a 1-to-1 match to numbers that are based on short presentations, they nonetheless provide a good range in which the respective individual should show variations in pre-dominance around an equi-dominance for both interpretations. For the continuous task, we therefore presented those 5 stimulus versions to participants that corresponded to their individual 20%, 35%, 50%, 65% and 80% fits from the onset task. These stimuli were presented in 11 trials, for 80 s at a time (Fig. 1d). The 50% stimulus was shown in the first, the sixth and the eleventh trial, the order of the other stimuli was random (one presentation each in trials 2 through 5, one in trials 7 through 10). Participants were asked to report switches of their perception by a button press, one button for switches towards the "from-below", another for switches towards the "from-above" state. Crucially, participants were asked to report on perceptual switches only if they arose, to encourage a passive viewing and experience of perceptual switches. Button assignment in each individual was consistent with the onset task. Four of the 11 trials (one per cube version, randomly picked), were followed by a replay phase, in which strongly disambiguated Necker Cube versions were presented to induce perceptual changes exogenously at the same time-points as reports in the previous experimental phase. Replay phases were excluded from the analysis of behavioral data, as were rare occasions of reported switches that were followed by a "switch" to the same state (i.e., the same button was pressed twice in succession). The latter were also removed from imaging analysis (see below).

All tasks were coded in Matlab (Matlab and Statistics Toolbox Release 2013a, TheMathWorks, Inc., Natick, Massachusetts, United States) and the Psychophysics Toolbox (Brainard and Vision, 1997; Pelli, 1997; Kleiner et al., 2007).

## 2.4. Image acquisition

Scans were acquired on a Siemens 3T Prisma at the Center Hospitalier Universitaire Vaudois. Once settled in the bore, we acquired first a localizer scan, followed by a gre-field mapping scan to generate voxel displacement maps (64 slices;  $3 \times 3 \times 2.5$  mm resolution; FOV 192 mm; slice TR/TE 700/4.92 ms; FA = 80°; Base resolution = 64 mm); and a B1 mapping scan to correct for magnetic field inhomogeneities ( $4 \times 4 \times 4$  mm resolution; FOV = 256 mm; slice TR/TE 200/39.1 ms; Base resolution 64 mm). We then alerted participants to the beginning of the task and functional acquisition. Parameters for our EPI sequence were as follows: 2D EPI, Multi-Echo sequence (3 echo times),  $3 \times 3 \times 2.5$  mm resolution, FOV = 192 mm; FA = 90°, slice TR = 80 ms; TE = (17.4; 35.2; 53 ms); base resolution 64 mm; 34 slices; volume TR = 2.72 s; parallel acceleration mode = GRAPPA, with an acceleration factor = 2. At the end of the experimental session, anatomical T1-weighted images were acquired with the following sequence parameters: MPRAGE,  $1 \times 1 \times 1$  mm

**Fig. 1.** Stimuli and tasks. A) Procedure of the financial uncertainty task (the card game). Participants place their bet (whether the first or the second card will be higher) before the first card is revealed. Decision variables change when cards are drawn, although the decision process is already completed. Participants respond regarding their outcome to ensure they remain vigilant to the task. One trial lasts about 25 s. B) The 21 different versions of the Necker cube used in the onset task, neutral cube in the middle ("0"). C) Procedure of the "onset" Necker cube task. A version of the Necker cube is presented for 3s and participants respond whether they have seen it from above or below in a subsequent 2s interval. D) Procedure of the "continuous" Necker cube task. In "rivalry" phases, participants respond to endogenous changes in the perceptual interpretation of the presented constant Necker cube (adjusted per individual to yield five defined levels of bias in the onset task); in "replay" phases, strongly disambiguated versions mimic the preceding sequence of perceptual switches by exogenous changes. E) Mathematical models of decision variables in the Card Game (left) and Necker Cube (right) Tasks. Both tasks were subject to the same mathematical models of risk and surprise. In the Card Game, risk was computed based on the probability of winning a gamble, given the value of Card 1 (x-axis) and the bet placed, while risk prediction error (surprise) is derived from the difference between the expected risk and actual risk at the outcome of Card 2. The monetary unit on the y-axis is 1 CHF. In the Necker Cube task, risk corresponds to the variance of the probability of viewing a given stimulus from above (x-axis), and surprise computed based on the resulting percept at switch. The two diagrams above show that uncertainty related-variables (risk, surprise) are unsigned and quadratic, whereas first order errors (RPE and perceptual prediction errors) are linear. Most notable is the correspondence of the model across the two tasks to the different paradigms, stimuli and outcomes. (RPE – reward prediction error; Risk PE – risk prediction error; PE perceptual error).



**Fig. 2.** Behavioral data of the Necker cube tasks. A) Onset task: Fraction of “from above” reports for each of the 21 cubes (Fig. 1c) were fitted by a cumulative Gaussian per individual (gray lines), and to the average (black symbols, mean  $\pm$  standard error of mean (sem) across subjects; black line: fit to aggregated data. B) Continuous task: predominance of the ‘from above’ percept (i.e., time the ‘from above’ percept was reported divided by the sum of times either ‘from above’ or ‘from below’ was reported). x-axis denotes individual adjusted ‘from above’ level according to onset task. Mean and sem across participants. C) Median duration of a dominance phase, mean and sem across observers, x-axis as in panel B.

resolution; FOV = 256 mm; slice TR/TE = 2 ms/2.39 ms; FA = 9°; base resolution = 256 mm).

## 2.5. Image preprocessing

Functional scans were preprocessed and analyzed using SPM 12. As we employed a multi-echo EPI sequence with 3 echo times, we first

summed the three volumes to obtain one scan per TR (Kettinger et al., 2016; Poser and Norris, 2009). We applied slice-timing correction to the volumes, as they were acquired with a comparatively long effective TR (2.72 s) (Sladky et al., 2011). We then generated voxel displacement maps (VDM) and applied these to functional volumes. Volumes were warped and realigned to the mean functional image using least squares and a 6 parameter (translations and rotations in space), rigid-body transformation to correct for intra-session motion artifacts, before applying a bias-field correction. Then individual T1-weighted volumes were co-registered to the mean functional image using a rigid body model, estimated with mutual information. The T1 image was then segmented (6 class tissue probability maps) and normalized to MNI space using unified segmentation (Ashburner and Friston, 2005). These normalization parameters were then applied to functional volumes. Volumes were then smoothed with a Gaussian kernel of 8 mm FWHM (Friston et al., 1995). The resulting images were used for analysis.

## 2.6. Mathematical models of uncertainty

We modeled first- and second-order uncertainty in the predictive (indexed by expected outcomes, and their variance) and in the outcome phase of a decision (formalized by errors in expected outcomes and errors in their expected variance). Variance has interchangeably been called risk in economic circles (Markowitz and March, 1952) and expected uncertainty in neuroeconomics studies (Payzan-Le Nestour et al., 2013). Errors arising from predicted uncertainty have been dubbed surprise (Preuschoff et al., 2006; Faraji et al., 2018). In our tasks, we employ mathematical models of uncertainty and surprise that have previously been used for the financial uncertainty task (Preuschoff et al., 2006, 2008a, 2011).

Risk is analogous to mathematical models of entropy (Quartz, 2009; Faraji et al., 2018). We use risk prediction error as a measure of surprise, which yields values that are highly correlated with alternative models of surprise, notably absolute prediction error (Preuschoff et al., 2011) and Shannon information (Strange et al., 2005) (see Appendix 3). Therefore we do not consider alternative accounts of predicted uncertainty or surprise as competing models to our study. We borrow the same accounts used in our previous studies and adapt them to our perceptual uncertainty task in order to build on previous results relating to economic decision-making (Fig. 1e).

Our measures of perceptual uncertainty are based on the variance of the probability distribution associated with viewing the cube from-above. Each Necker Cube stimulus is endowed with an expected value, the average of the probability of viewing a cube from-above and the probability of viewing the cube from-below. We derive 2 measures of perceptual uncertainty associated with a Necker Cube stimulus, one objective and one subjective. For subjective uncertainty, we compute the squared difference of a probability associated with viewing a given stimulus from-above from the expected value of that same stimulus squared, based on individual responses in the Necker Cube – onset task. Objective uncertainty is derived by computing the same formula, but with assumed probabilities of viewing the cube from above based on the cube’s warp bias. These measures mirror risk accounts used in economic contexts but for one difference; expected values in the latter are associated with a monetary gain, or reward. Here, we substitute a monetary gain value of +1 with a from-above perceptual outcome of +1; similarly, we substitute a monetary loss value of -1 with a from-below perceptual outcome of -1. Finally, we derive a prediction error related to the switch by subtracting the stimulus’ expected value from the perceptual state following the perceptual switch; and a risk prediction error captured by the difference between the squared prediction error and the stimulus-associated risk value.

Equation (1): Predicted Reward

$$E[\text{reward}] = \Sigma p(\text{outcome} | \text{card } 1, \text{ bet}) \quad (1)$$

Equation (2): Reward Prediction Error

$$\delta = \text{Reward} - E[\text{Reward}] \quad (2)$$

Equation (3): Risk Prediction Error (Surprise)

$$\begin{aligned} Ripe_{win} &= 4 * p_{win} + 8 * p_{win}^2 \\ Ripe_{loss} &= 4 - 12 * p_{win} + 8 * p_{win}^2 \end{aligned} \quad (3)$$

Equation (4) Expected Percept

$$E[\text{Percept}] = (p_{from-above} * 1) + (p_{from-below} * -1) \quad (4)$$

Equation (5) Perceptual Uncertainty (Perceptual Risk)

$$PRisk = p_{from-above} * (1 - E[\text{Percept}])^2 + p_{from-below} * (-1 - E[\text{Percept}])^2 \quad (5)$$

Equation (6) Surprise As Risk Prediction Error for Perceptual Switches

$$\begin{aligned} PRipe_{from-above} &= 4 * p_{from-above} + 8 * p_{from-above}^2 \\ PRipe_{from-below} &= 4 - 12 * p_{from-above} + 8 * p_{from-above}^2 \end{aligned} \quad (6)$$

## 2.7. Correlations between decision variables

The expected values of a percept, respectively, reward, are distinct from expected uncertainty, as the first represents a linear function while the second is quadratic. The same relationship can be applied to first and second order prediction errors. We pooled all emergent decision variables across trials in all sessions and participants in the Financial Uncertainty task (see Table 1).

Similarly, first and second-order perceptual prediction errors across all participants showed little correlation ( $R^2 = 0.1121$ ).

## 2.8. fMRI data analysis

Functional data analysis was implemented in SPM12b ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)). We performed a model-based fMRI analysis on all 3 experimental tasks (Card game, Necker onset, Necker continuous) using a summary statistic approach (Worsley et al., 2002). To address our research question, we first analyzed individual subject EPI time-series using a fixed-effects general linear model (GLM) before pooling subject-level contrast images in a random effects analysis, where contrasts of interest were examined with one-sample t-tests to obtain an average response. Regressors in the respective GLMs were convolved with the hemodynamic response function (Glover, 1999); an AR(1) model was employed to address autocorrelation in the timeseries.

Regressors for specific onsets were parametrically modulated by decision variables (see Fig. 2). We applied the same computational models to both the financial and perceptual uncertainty paradigms via parametric modulation. Each parametric modulator was serially orthogonalized with respect to the event onset and sequentially to prior parametric modulators. While our main research question rests on a strong *a priori* hypothesis concerning a target brain region (anterior

**Table 1**

Correlation coefficients between decision variables in the financial uncertainty task.

	Expected Reward	Expected Uncertainty	Reward Prediction Error	Surprise
Expected Reward	-	-0.03	0.00	-0.01
Expected Uncertainty	-	-	-0.02	0
Reward Prediction Error	-	-	-	0

insula) (Platt and Huettel, 2008; Grinband et al., 2006; Brevers et al., 2015; Preuschoff et al., 2008a; Mohr et al., 2010; Rutledge et al., 2010), and we also do not expect dramatic signal changes, we nonetheless opted to conduct whole-brain analyses for all univariate models below, with the aim of 1) highlighting the specificity of the anterior insula with respect to uncertainty; 2) conforming to best practices relating to fMRI analyses (Eklund et al., 2016). Below, we describe GLMs applied to each task in greater detail.

### 2.8.1. Financial Uncertainty

In the card game, we modeled the 5.5 s period after the participant hears the value of card 1 as a boxcar function and further parametrically modulated these events with expected reward followed by a risk value; and we further modeled the 5.5 s period after the participants hears the value of card 2 with a boxcar function and parametrically modulated these epochs with a reward prediction error followed by a risk prediction error. The model also included two nuisance regressors including events modeled as delta functions, including one for button press events (0s duration), to report the bet and the trial outcome, and one for sound, upon hearing instructions for the bet placement and the trial report (0s duration) (Fig. 2c). Expected reward and risk, as well as reward prediction error and risk prediction error are not expected to correlate (Fig. 1e) and show little evidence of multicollinearity in subject-level design matrices. These quantities are further subject to serial orthogonalization in the SPM software. Also included in the GLM were 6 movement related regressors of no interest.

### 2.8.2. Necker Cube - onset task

We constructed 2 general linear models containing 2 experimental regressors as well as 6 movement-related regressors of no interest. The GLMs differed along one dimension; specifically, in one GLM, perceptual risk was computed based on objective probabilities of viewing a trial-related stimulus from-above, or probabilities related to the degree of experimentally manipulated cube bias towards one percept or another; while in the other, perceptual risk values were derived from subjective probabilities of perceiving one cube representation or another, computed at the end of the task from individual participants' reports. The first regressor in both GLMs contained onsets for stimulus presentation, modeled as boxcar functions of 3 s (the viewing period), parametrically modulated by risk values. The second regressor in the GLM contained onsets for button presses during the response window modeled as delta functions (Fig. 2a).

### 2.8.3. Perceptual surprise, Necker Cube continuous task

In the continuous task of the perceptual uncertainty experiment, we constructed a GLM with the following regressors. Our primary regressor of interest was composed of perceptual switches modeled as delta functions, parametrically modulated by the perceptual prediction error, followed by the risk prediction error, or surprise, and finally by the amount of time spent in the previous perceptual state, to control for potential time-related effects. We eschewed the analysis of risk prediction in this task because such an analysis presents 2 problems: the onset of the dominance times are the same as the those of the switches, introducing a problem of multi-collinearity; second, dominance times vary widely within and across individuals. Investigating onsets with variable durations parametrically modulated by risk suggests that the trial-by-trial risk variability may also be modulated in strength by time alone, which may confound probing of perceptual risk. For these reasons, we confined the study of perceptual risk to the onset task.

The risk prediction error was computed based on the participant's individual probability of seeing a stimulus from-above or from-below and modeled each switch by this value given both the stimulus and the perceptual state into which the participants switched to. The 5 different levels of cube warping towards a from-above state (20%, 35%, 50%, 65% and 80%) yielded 5 values of surprise, with maximal surprise emerging for least likely perceptual states in most disambiguated stimuli (e.g.,

switching into a “from-below” perceptual state for a cube biased at 80%). We further included replay condition presses as a regressor in the model to account for BOLD responses to exogenous stimulus switches. These regressors were convolved with the canonical HRF. We also included 6 motion-related regressors of no interest in the GLM (Fig. 2b).

### 3. Results

#### 3.1. Financial uncertainty task - behavioral results

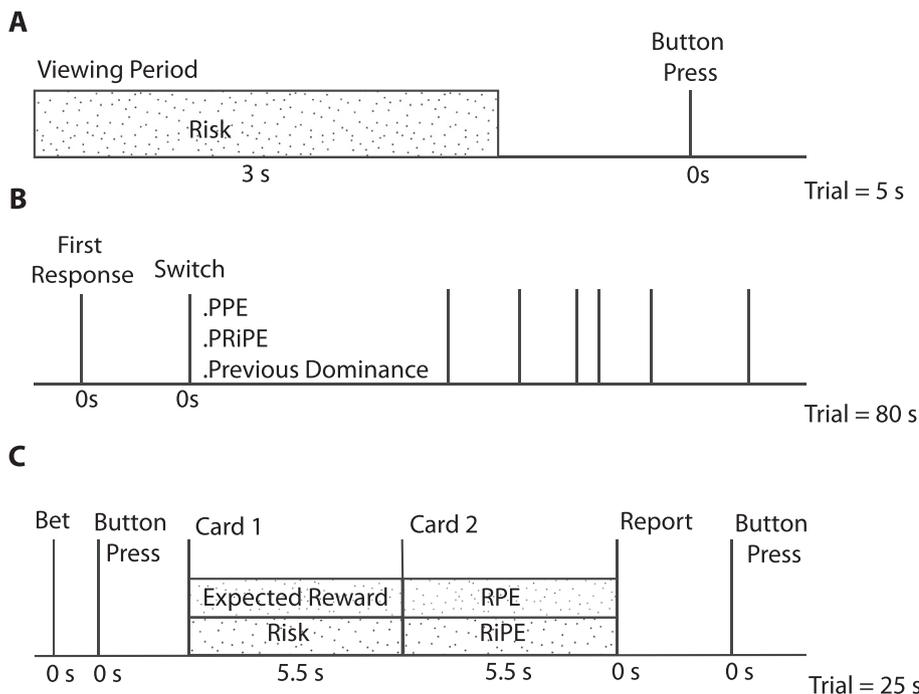
Of 29 participants, four were excluded for technical reasons. Twenty-five remaining participants were included in the analysis of the financial uncertainty task. Participants were paid for one out of the three blocks performed. Average task-related payout per participant was 29.57 CHF; across all blocks and participants, payoffs were in the range of 13–39 CHF. Trials where participants either missed a bet placement or provided an incorrect response to the trial’s outcome were excluded from analysis. Participants who showed an overall error rate of >30% were excluded from analysis (1 participant). Error rates ranged from 0 to 26.67%, with the majority of participants (15) showing error rates of 1.11%.

We tested for mean differences between higher and lower bets across sessions and found no significant difference ( $F = 0.19, p = 0.8324, df = 2$ ). There was a significant difference in bet choice in all tasks however, with participants choosing a higher over a lower bet in all sessions ( $F = 34.69, p < 0.001, df = 1$ ). We cannot conclude that there is a bias towards selecting higher bets in spite of this result, because buttons indicating a higher bet were consistently on the right side for all subjects. Significant results may reflect handedness rather than bet preference. We further examined “strategic” bets within sessions, by tallying bet switches following a loss with bet maintenance following a win, as a measure of previous-bet influence, compared to non-strategic counterparts. We find a significant preference ( $t = -3.01, p = 0.0035, df = 74$ ) for non-strategic bet placement, suggesting participants understood the random nature of the task and did not deploy a strategy. However, we also performed binomial tests on the proportion of strategic to non-strategic bets for each session, tested against the null (50%) and found 14 sessions to show a significant preference for “strategic” betting.

#### 3.2. Perceptual uncertainty tasks - behavioral results

In the onset task, participants were briefly presented 21 different versions of the Necker Cube and the percentage of “view from-above” was determined across 10 presentations of each. One participant, who classified all versions of the Necker as seen from-above, was excluded from further analysis. As intended, “from-above” percepts increased with warping level and were well fitted by a cumulative Gaussian function (Fig. 3a, gray). The “point of subjective equality” – i.e., the warping at which 50% “from-above” was perceived - was slightly shifted towards warpings indicating “from-below” (Fig. 3a). This is consistent with a “from-above” bias for the symmetric configuration. The fitting of the Necker Cube onset responses yielded an average  $R^2 = 0.80$ ; and  $k = 4.07$  across 22 subjects.

From the individual fits, we determined the cube versions at which the individual had 20%, 35%, 50%, 65% and 80% dominance for the “view from-above” at onset. The corresponding cubes were used for the continuous task in the respective individual. From the analysis of this task, one additional participant who on average reported less than 1 perceptual transition per condition, was excluded. Since the task differed between the onset task (respond to a cube onset) and the continuous task (report switches during continuous viewing) it could not be expected that predominance of the from-above percept (i.e., the percentage of time the from-above percept is seen as dominant) would match 1-to-1 to the nominal values of 20% ... 80%. Nonetheless, we found that the manipulation was successful: in the rivalry conditions of the continuous task, there was a significant main effect of manipulation on predominance ( $F(4,80) = 22.8, p < 0.001$ ; repeated-measures ANOVA) and the increase with the manipulation level was monotonic and as intended (Fig. 3b). Importantly, there was also a significant main effect of the cube manipulation on dominance durations ( $F(4,80) = 3.53, p = 0.01$ ) with a minimum at the intermediate manipulation (“50%”) and increasing values towards the more extreme cubes (Fig. 3c). This is consistent with the intended experimental manipulation: high levels of uncertainty at the “50%” level, medium levels of uncertainty at the “35%” and “65%” level and low levels of uncertainty at the “20%” and “80%” level. We fit dominance times from all individuals to a gamma distribution to



**Fig. 3.** Graphical representation of general linear models for the three tasks. A) The Necker Cube Onset task GLM was comprised of a regressor for the viewing period of the cube (3s) parametrically modulated by stimulus risk, as well as a regressor including button responses. B) The Necker Cube Continuous Task included regressors for first response onsets, and a regressor for switch responses, parametrically modulated by first order perceptual prediction error; perceptual risk prediction error; and previous dominance time. C) The Card Game GLM included two regressors of no interest comprised of button presses (for bet and trial outcome), modeled as stick functions; and of sound (for request to bet and to report trial outcome), also modeled as stick functions. The sounding of cards 1 and 2 were modeled with boxcar functions (5.5s durations), parametrically modulated by expected reward and risk; and Reward Prediction and Risk Prediction Error, respectively. (PPE: perceptual prediction error; PRiPE: perceptual risk prediction error; RPE: reward prediction error; RiPE: risk prediction error).

determine the temporal characteristics of perceptual switches (Zhou et al., 2004) and find that Necker Cubes across all bias values conform to such a distribution (Figure A2; for individual fits, see Appendix 4).

Participants displayed significant variability in number of switches reported, with a mean rate of 115 switches (SD = 67, range (17, 266)).

### 3.3. Imaging results

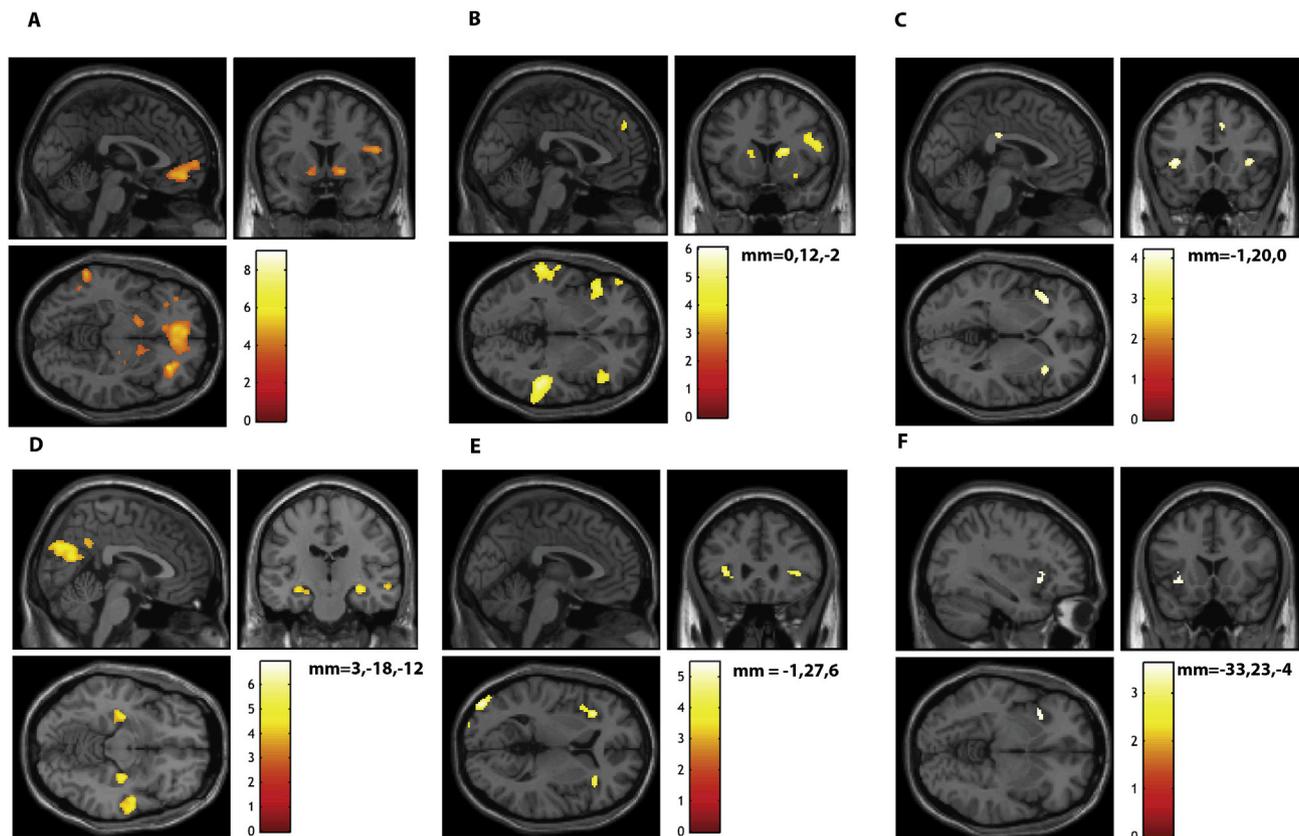
Functional data analysis was implemented in SPM12b ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm)) at the single subject level, and the SnPM toolbox (<http://warwick.ac.uk/snpm>) for group level analyses. Our second level analysis comprised whole-brain permutation tests (10 000 permutations), with a variance smoothing of 8 mm. We applied a false discovery rate (FDR; Benjamini and Hochberg, 1995) threshold of 0.05 at the voxel-level to control for false discovery rates (Han and Glenn, 2018). For completeness, we also report results using a whole-brain family-wise error corrected threshold of  $p = 0.05$  at the voxel level, to control for Type 1 errors (Eklund et al., 2016). Cluster sizes below are reported in voxels; left and right designate neurological orientation.

#### 3.3.1. Financial Uncertainty

3.3.1.1. Contrast reward prediction error (RPE). We performed a one-

sample  $t$ -test on the parametrically modulated outcome of each trial by its corresponding RPE value. A positive signed  $t$ -test on this regressor using FDR correction yields significant clusters in expected areas, notably bilaterally in the caudate (right and left,  $p = 0.009$  and  $0.015$ ,  $t = 4.26$  and  $3.95$ ,  $k = 250$  and  $361$ ) and the cingulate cortex (left middle and right posterior,  $p = 0.017$  and  $0.022$ ,  $t = 3.93$  and  $3.87$ ,  $k = 74$  and  $67$ ). We further find significant clusters bilaterally in the inferior frontal gyrus ( $p = 0.009$  and  $0.009$ ,  $t = 6.94$  and  $4.86$ ,  $k = 7910$  and  $1012$ ), left hippocampus ( $p = 0.009$ ,  $k = 75$ ,  $t = 3.72$ ); and right middle temporal gyrus ( $p = 0.011$ ,  $t = 4.21$  and  $k = 114$ ) (Fig. 4a). When controlling for FWE, we find significant peaks in right inferior frontal gyrus ( $p = 0.001$ ;  $t = 6.94$ ;  $k = 150$ ); left medial frontal cortex ( $p = 0.004$ ;  $t = 5.57$ ;  $k = 202$ ); left superior frontal gyrus ( $p = 0.022$ ;  $t = 4.92$ ;  $k = 10$ ); left middle temporal gyrus ( $p = 0.036$ ;  $t = 4.74$ ;  $k = 5$ ).

3.3.1.2. Contrast risk prediction error (RiPE). Trial outcomes at Card 2 were further modulated by their risk prediction errors (surprise) in addition to their reward prediction errors. We replicate previously reported results for surprise at Card 2, where we find significant clusters in the left and right insulae ( $p = 0.031$  and  $0.031$ ,  $t = 4.70$  and  $4.56$ ,  $k = 870$  and  $1597$ ) and the right caudate ( $p = 0.029$ ,  $t = 4.96$ ,  $k = 293$ ), as well as left and right medial temporal gyri ( $p = 0.029$  and  $0.029$ ,  $t = 6.11$  and  $5.33$ ,  $k = 1237$  and  $694$ ) (Fig. 4d). While memory regions are not of



**Fig. 4.** A) Contrast for reward prediction error in the Card Game at the gamble's outcome shows an expected pattern of BOLD responses in the dorsal striatum (bilateral caudate) and cingulate. B) Surprise contrasts in the financial uncertainty paradigm find significant BOLD responses in the bilateral anterior insula, and middle temporal gyri and right caudate, replicating previous findings. Shown are non-parametric statistical maps thresholded at  $p = 0.05$ , FDR corrected. C) fMRI results for Perceptual Risk in Necker Cube Onset Task. Risk modulation in perceptual uncertainty yields significant clusters in the bilateral insula. The viewing period of the Necker Cube (3 s) was modulated by a measure of subjective risk, based on the probability of a participant viewing a particular Necker Cube stimulus from above or from below. Images show non-parametric statistical maps thresholded at  $p = 0.05$ , FDR corrected. D) Perceptual prediction errors in the Necker Cube Continuous task show significant clusters in bilateral. Hippocampus, and the cuneus and also yielded several significant clusters in visual areas. E) The contrast for surprise modeled at the perceptual switch response in the Necker Cube Continuous task yields significant BOLD responses in the anterior insula as well as in occipital regions. Shown are non-parametric statistical maps thresholded at  $p = 0.05$ , FDR corrected. F) Thresholded maps are shown for the conjunction analyses of surprise contrasts in the financial and perceptual tasks, masked for the left anterior insula only. These contrasts were tested against the conjunction null and the analysis constrained to one anatomical map of both left and right anterior insulae. Only the left anterior insula cluster survived FWE correction for multiple comparisons. All colorbars above represent  $t$ -values.

direct interest to our research question, significant clusters in the medial temporal gyri support the notion that surprise is predicated on comparing current evidence against past predictions. Controlling the FWE rate, we find significant responses in left and right middle temporal gyrus ( $p < 0.001$  and  $p = 0.006$ ;  $t = 6.11$  and  $t = 5.33$ ;  $k = 118$  and  $k = 32$ ); right caudate ( $p = 0.019$ ;  $t = 4.96$ ;  $k = 6$ ); and left anterior insula ( $p = 0.0429$ ;  $t = 4.7$ ;  $k = 4$ ).

### 3.3.2. Perceptual Uncertainty

**3.3.2.1. Onset task – contrast risk.** We performed a *t*-test on the viewing period of the cube (3s) prior to the participant's percept selection, modulated by the stimulus risk to probe perceptual risk activation. Perceptual risk was defined as the variance related to the average perceptual value of a stimulus, based on the probabilities of its being viewed from-above or from-below. Using experimentally derived subjective risk values, whole brain analysis reveals responses in the left and right anterior insula ( $p = 0.009$ ,  $t = 4.10$ ,  $k = 74$ ;  $p = 0.014$ ,  $t = 3.95$ ,  $k = 44$ ) (Fig. 4c). Objective risk values yielded no significant voxels, underlining the importance of accounting for individual differences in perception. Using a FWE error rate on the same data, we find the same significant clusters in left and right insula ( $p = 0.009$  and  $p = 0.014$ ;  $t = 4.10$  and  $t = 3.95$ ;  $k = 74$  and  $k = 44$ ).

**3.3.2.2. Continuous task - contrast perceptual prediction error.** In the Necker Cube Continuous task, we modulated the manual response to a perceptual switch with 3 values: the first-order perceptual prediction error; the time spent in the previous perceptual state, to account for BOLD responses related to dominance times; and a perceptual risk prediction error. We first performed a contrast on the perceptual prediction error, to account for a first-order error as we do in the Card Game for reward prediction error. Results show significant clusters in the precuneus ( $p = 0.035$ ,  $t = 5.20$ ,  $k = 2231$ ), left and right angular gyri ( $p = 0.035$  and  $0.046$ ,  $t = 4.75$  and  $4.34$ ,  $k = 882$  and  $19$ ), left and right occipital poles ( $p = 0.040$  and  $0.035$ ,  $t = 4.24$  and  $4.41$ ,  $k = 32$  and  $68$ ), left and right hippocampi ( $p = 0.035$  and  $0.035$ ,  $t = 4.34$  and  $4.56$ ,  $k = 232$  and  $122$ ), and right middle temporal gyri ( $p = 0.035$ ,  $t = 4.75$ ,  $k = 392$ ) for this contrast (Fig. 4b). Thus the pattern here shows a response in memory regions (hippocampus and middle temporal gyrus) in addition to the visual system, suggesting the error calls on higher-level functions such as memory, to compare previously held expectations to the current evidence, in addition to perceptual processes. Applying a FWE correction yields significant results in the left precuneus ( $p = 0.024$ ;  $t = 5.20$ ;  $k = 31$ ).

**3.3.2.3. Continuous task - contrast perceptual risk prediction error (surprise).** A positive contrast for the perceptual risk prediction error yielded significant clusters in several regions of the visual system, including bilateral mid, inferior and fusiform occipital gyri and left calcarine cortex. In addition however, significant BOLD responses were found bilaterally in the anterior insulae ( $p = 0.035$  and  $0.040$ ,  $t = 4.34$  and  $3.85$ ,  $k = 172$  and  $56$ ) (Fig. 4e). Using FWE, we find significant responses in the left middle occipital gyrus ( $p = 0.014$ ;  $t = 5.34$ ;  $k = 38$ ); right precentral gyrus ( $p = 0.025$ ;  $t = 5.07$ ;  $k = 8$ ); right occipital fusiform gyrus ( $p = 0.029$ ;  $t = 5.01$ ;  $k = 12$ ); and left inferior occipital gyrus ( $p = 0.030$ ;  $t = 4.99$ ;  $k = 8$ ).

We further examined whether surprise values associated with exogenous switches in the replay condition contributed to anterior insular response. We designed a GLM including a regressor for both report and replay onsets, parametrically modulated by their first and second order prediction errors. Stimulus switch in the replay condition is expected and therefore should not incur surprise, however there is a possibility that a stimulus change, be it exogenous or endogenous, drives anterior insular response. Average coefficient values across participants in the anterior insula for both report and replay surprise are decreased relative to report

only, although not significantly so ( $t = 1.33$ ,  $p = 0.2$ ,  $df = 20$ ) (Appendix 1).

### 3.3.3. Financial and perceptual risk prediction error (surprise) – conjunction analysis

To test for a common BOLD response to surprise across the two tasks, we performed a conjunction analysis of the surprise contrasts in the financial and perceptual paradigms ( $N = 25$  and  $N = 21$ , respectively). We performed this analysis in SPM by selecting a two-sample *t*-test design for non-independent samples of unequal variance. We performed this test at the whole brain, peak level using a FWE error corrected threshold. We tested against the conjunction null (logical AND), to determine which voxels were found significant in both tasks (Friston et al., 2005) and found a significant cluster in the left anterior insula ( $p = 0.042$ , FWE corrected;  $k = 33$ ; peak voxel coordinates  $-36, 16, -4$ ) as well as a non-significant cluster in the right anterior insula ( $p = 0.078$ , FWE corrected;  $k = 8$ ; peak voxel coordinates  $34, 24, -4$ ) (Fig. 4f). FDR correction could not be used in this instance, because all significant voxels were clustered in the anterior insula exclusively; this imposes the false discovery rate on said clusters, yielding an infinite T threshold of significance.

### 3.3.4. Axiomatic model testing of the risk prediction error in the anterior insula

Reward prediction errors in the striatum have been shown to act as true prediction errors (Rutledge et al., 2010), meaning regional striatal responses track with increasing RPE. Although a corollary linear relationship of risk prediction error with the anterior insula is implied in the model-based analysis above, it could be driven by one particular high level of surprise included in the task design. To determine if the anterior insula tracks increasing levels of surprise, we designed a general linear model where card game outcome and perceptual switch onsets were categorized into three levels of respective first order prediction error. We extracted individual mean anterior insula values for each of these first-level contrasts and averaged these for each level of surprise, for each task. A true risk prediction error encoded in the anterior insula should yield a v-shaped curve for each task, with lowest and highest first order PE both yielding high surprise. We find this pattern for both tasks, highlighting that the anterior insula encodes surprise across both task domains (Fig. 5), although the two curves neither overlap perfectly nor are they symmetrical. We performed an ANOVA with surprise level and task as factors. We find a significant effect of surprise levels on anterior insula mean coefficient values ( $F = 6.63$ ,  $p = 0.018$ ,  $df = 2$ ) but no effect of task ( $F = 0.22$ ,  $p = 0.6403$ ,  $df = 1$ ).

### 3.3.5. Decoding surprise levels in perception from surprise levels in financial tasks

We then performed a classification of beta values extracted from a pre-defined set of regions to support our results in the model-based univariate analysis above. Our aim was to determine if surprise-related neural responses in Financial Uncertainty could accurately predict those in Perceptual Uncertainty.

To address this question, we constructed two general linear models, one for the Card Game and one for the Necker Cube Continuous Task where Card 2 onsets, respectively perceptual switches, were separated into low, medium and high levels of surprise. For the Card Game, low surprise included the lowest 3 values; mid surprise, the following 3 surprise values in the ordinal scale; and for high surprise, the highest 3 values of surprise. For the perceptual task, low surprise included switches for the fully ambiguous Necker Cube; mid surprise, the 35%/65% Necker cubes; and high surprise, the 20%/80% Necker cubes.

Our study was based on previous work identifying the striatum and the anterior insula as neural correlates to uncertainty. Our *a priori* list of candidate regions for surprise therefore included bilateral anterior insulae; caudate; putamen; and ventral striatum, totaling 8 features for our classification in all.

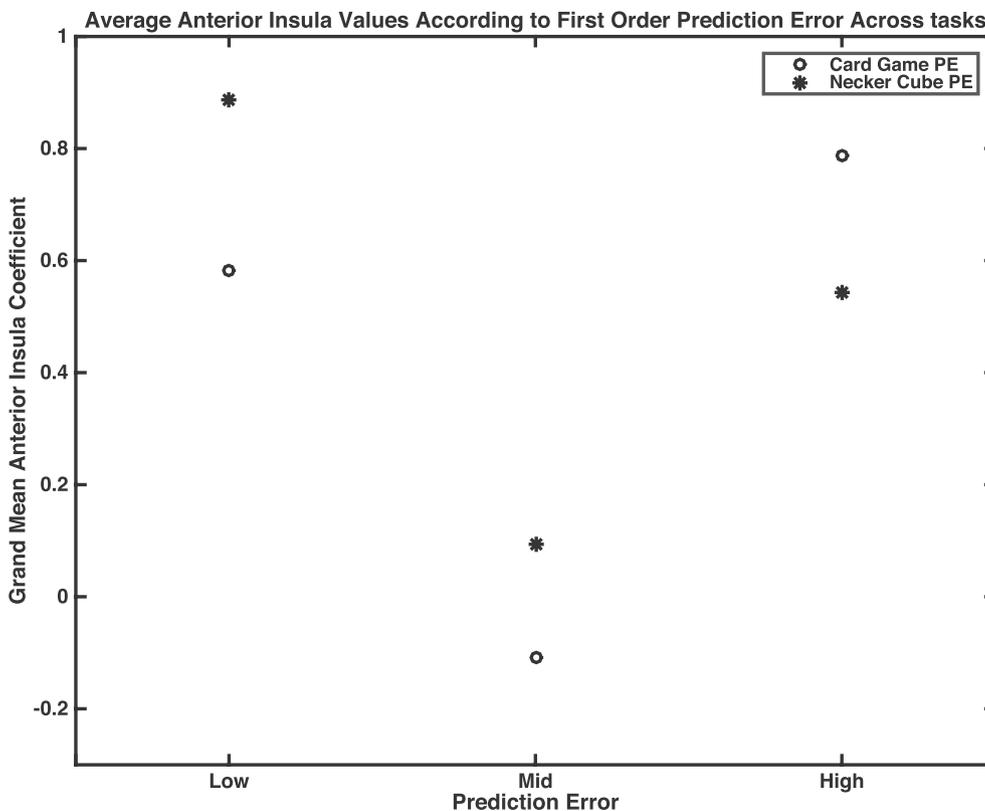


Fig. 5. Tracking surprise as a prediction error in the Card Game and Necker Cube Continuous Tasks. Onsets of Card 2's sounding and perceptual switches, respectively, were separated according to first order prediction level (low, medium and high). A first level analysis was then performed on each surprise level for each subject. Coefficient estimates for each contrast in the anterior insula (left and right summed masks) were extracted, averaged, and plotted according to surprise level. Both low and high first order prediction errors should incur high surprise in the anterior insula. First order prediction errors incur just such a pattern, as shown in the graph above.

Our training set was composed of beta values for the above regions corresponding to low, medium and high surprise in the Card Game for all participants included in the Financial Uncertainty experiment and our test set, regional beta values for all participants included in the Necker Cube continuous task. Beta values represented predictors, while surprise labels comprised outcome variables. We performed an initial feature selection on the regions as a means of dimension reduction using an ANOVA, and retaining only those regions whose classification score was significant at the  $p = 0.05$  threshold. The following regions were retained: left anterior insula ( $F = 6.18$ ;  $p = 0.003$ ); right anterior insula ( $F = 4.78$ ;  $p = 0.011$ ); left caudate ( $F = 3.24$ ,  $p = 0.045$ ); and left putamen ( $F = 3.20$ ,  $p = 0.047$ ).

We employed a random forest classifier to predict perceptual surprise from financial Surprise and found a classification accuracy of 38%. Of the four features retained to perform this classification, the right anterior insula, followed by the left anterior insula, were the regions that best contributed to accurate predictions. Above-chance accuracy is often taken as a measure of predictive power in machine learning paradigms, but caution must be taken in the case of our results, as an accuracy of 38% presents low predictive power (Combrisson and Jerbi, 2015). We performed a permutation test to examine the significance of the accuracy score above (1000 permutations) and obtained a classification score of 0.45,  $p = 0.004$ . It must be noted that accuracy for low and mid surprise classification exceeded that of high surprise in particular (F1 scores of 0.45, 0.48 and 0.09 for low, mid and high surprise, respectively).

#### 4. Discussion

To probe the presence of a generalized neural correlate of the inferential process in both value-based decision-making and perception, we applied formal accounts of risk and surprise (Preuschoff et al., 2008b) to BOLD signals arising from uncertainty in financial and perceptual domains. Using the Necker Cube to elicit perceptual uncertainty, behavioral results show that experimental manipulation of the stimulus succeeded in expanding the range of probabilities associated with two percepts,

allowing for a derivation of perceptual uncertainty. To capture the endogenous inferential process related to perception, we model fMRI BOLD responses to viewing ambiguous stimuli with a formal account of predicted uncertainty (risk), finding a significant response in the anterior insula. We then modeled perceptual switches with a formal account of predicted uncertainty error (surprise), and found a response in that same region. We further administered a well-established gambling task in the same individuals, replicating results in the anterior insula for surprise. Insular responses to surprise across tasks were found only when applying a false discovery rate, (Benjamin & Hochberg) and did not survive the more conservative FWE correction (Han and Glenn, 2018). A whole-brain conjunction analysis of the two tasks nonetheless finds a significant cluster exclusively in the left anterior insula. A decoding analysis further finds anterior insula response to financial surprise predicts regional responses to perceptual surprise. Our results suggest a common neural system dedicated to processing uncertainty in the anterior insula and more generally, support the theoretical framework of the brain as an inference machine (Friston, 2010), irrespective of stimulus or goal features.

Our tasks differed in several dimensions, retaining a commonality in uncertainty and hence, an unavoidable inferential process. We lift an objective, explicit model of uncertainty from economics, and apply it to a perceptual paradigm, with all that the latter implies: spontaneity, subjectivity and absence of conscious deliberation. In the Card Game, surprise arises in response to a changing external stimulus (card value), while in the Necker Cube surprise emerges from an internal change to a constant external stimulus. The distinction between internal and external attribution of uncertainty (Kahneman and Tversky, 1982) is by-passed with the use of both a computational framework and neuroimaging. Cross-domain studies of decision-making are rare (Basten et al., 2010) but we ventured across the sub-disciplines of neuroeconomics and psychophysics, motivated by two assumptions: first, the historical Helmholtzian view of inferential processes in relation to brain function, and second, the notion that computational accounts of decision-making, if true, should hold for any decision type (O'Connell et al., 2018).

#### 4.1. Perceptual switches as involuntary decisions

Several fMRI studies have probed perceptual switches (Frässle et al., 2014; Sterzer et al., 2009), but we examine their underlying, latent decision variables with a computational framework applied to fMRI signals, which allows us to empirically capture the hidden process of perceptual uncertainty. The Necker Cube provokes sharp transitions between perceptual states, providing an ideal stimulus to capture perceptual error, unlike binocular rivalry for instance, where more gradual transitions exist, introducing a temporal ambiguity in relation to the emergence of the error (e.g., Leopold et al., 1998; Naber et al., 2011). Perceptual switches are thought to be spontaneous and involuntary (Sterzer and Kleinschmidt, 2007), but there are questions on whether they are subject to volitional control (Hugrass and Crewther, 2012; van Ee et al., 2005). In the study above, participants were encouraged to view stimuli passively, to control for willful perceptual switches. Further, dominance times recorded follow a stochastic time-course, suggesting spontaneous switches in perception.

#### 4.2. Perceptual switches as prediction errors

By casting perceptual switches as prediction errors, we allay controversies regarding top-down (Wang et al., 2013; Long and Toppino, 2004; Sterzer et al., 2009) versus bottom-up (Polonsky et al., 2000; Parkkonen et al., 2008; Pearson et al., 2007) processing in ambiguous stimuli, as the iterative process of prediction and error exchanges information between low- and high-level areas (Rao and Ballard, 1999; Summerfield and De Lange, 2014). Nonetheless, few studies have explicitly modeled switches as prediction errors. Sundaeswara and Schrater (2008) characterize switches as inferential processes by applying a Markov Renewal Process to Necker Cube dominance times. In a fMRI study, a Bayesian account modeled switches in Lissajous figures as prediction errors (Weinhammer et al., 2017). While both stimulus and formal account differed from ours, Weinhammer and colleagues nonetheless find a response in the anterior insula to perceptual switches, in line with our results.

#### 4.3. Parsimony in model choice

The brain's frugality (Bullmore and Sporns, 2012) suggests it deploys parsimonious computational models in resolving uncertainty (Schwarzenbeck et al., 2015; Friston et al., 2012a,b; Clark, 2017). Various models have been used to quantify uncertainty (Rao, 2010; Kepecs and Mainen, 2012; Vilares et al., 2012) but we opted to couch the mean-variance theorem (accounting for risk), within a predictive coding framework (to measure surprise), first to build on our previous work (Preuschoff et al., 2006, 2008a, 2011) but also due to its simplicity. In economics, the mean-variance framework encapsulates uncertainty by integrating both the mean and variance (risk) of a utility distribution (Kroll et al., 1984; Markowitz and March, 1952). From a homeostatic perspective, this framework presents an ideal model of uncertainty because decision-making with only two values leaves an agent neurally unencumbered, and accounting for risk facilitates learning, as estimating an option's range shortens the trial-and-error process (d'Acremont and Bossaerts, 2008). It is worthwhile to note that several models can capture uncertainty (Friston, 2010) and some may be better suited to alternative paradigms, such as ones with an explicit learning component.

#### 4.4. Neural specificity of first order prediction errors

We replicate previous neuroimaging results for reward prediction error, notably its striatal correlate. Perceptual errors however correlate with a distinct response pattern in visual areas, suggesting these first-order errors elicit feature-specific neural responses; these errors nonetheless also implicate the hippocampus and temporal lobe, suggesting they recruit activity at the cognitive level, as would be expected in predictive coding (Clark, 2013). In contrast to reinforcement learning

studies (Schultz, 2010; Kuhnen and Knutson, 2005; Carlson et al., 2011), we do not find a striatal involvement for perceptual error. This absence of striatal response in the perceptual domain supports a neural specificity to first-order prediction errors (Clark, 2013). The striatum has been hypothesized to act as a learning center and not just a reward hub (Balleine et al., 2007; Hare et al., 2008; Tricomi et al., 2009; Daniel and Pollmann, 2014), an assumption that would see striatal responses to any prediction errors, including perceptual ones. Our results may challenge this postulated role for the region but could also indicate the absence of overt learning in our task. First-order prediction errors may account for case-specific decision features: BOLD responses in visual areas to perceptual prediction errors may reflect upstream, domain-specific, low-level errors (Parkkonen et al., 2008), while the insula may act as a common, downstream hub in the inference process.

#### 4.5. The insula in perceptual errors

Our model-based analyses on perceptual uncertainty and surprise implicate the anterior insula, in line with previous studies on value-based decision-making (Preuschoff et al., 2006; Payzan-Le Nestour et al., 2013; Platt and Huettel, 2008) and perception (Sterzer and Kleinschmidt, 2010). Several studies investigating perceptual decision-making find increased reaction time correlated with the anterior insula, which corroborates the hypothesis that decision uncertainty relates to the region specifically (Ho et al., 2009; Binder et al., 2004; Thielscher & Pessoa, 2010). Sterzer & Kleinschmidt's review on insular involvement in perceptual switches, titled "often observed but barely understood", suggests salience elicits a consistent response in the region but bypasses its plausible role in inferential processes (Singer et al., 2009). We propose that the insula responds to ambiguous perception because it is tuned to a conscious uncertainty. This insight is important because a formal framework can allow us to test populations that are less susceptible to illusions, as in autism (Pellicano and Burr, 2012) and schizophrenia (Schmack et al., 2015). Dampened insular responses to uncertainty and surprise in such patient populations may underlie observed deficits in decision-making.

#### 4.6. The insula in inference

A considerable body of evidence implicates insular BOLD responses in a wide range of functions including language (Ackermann and Riecker, 2010; Ardila et al., 2014); auditory processing (Bamiou et al., 2003; Bamiou et al., 2003); pain (Peyron et al., 2000; Corradi-Dell'Acqua et al., 2011); disgust (Wicker et al., 2003); gustatory function (Small, 2010); perception (Sterzer and Kleinschmidt, 2010); decision-making (Weller et al., 2009; Preuschoff et al., 2008a,b; Singer et al., 2009; Volz et al., 2005); uncertainty (Critchley et al., 2004; Xue et al., 2010; Jones et al., 2010; Weller et al., 2009); and emotion in general (Gasquoine, 2014). To account for all these functions, the insula has been cast as an interoceptive center, integrating bodily states into awareness (Craig and Craig, 2009). Interoception presents a specific, if broad, category that readily explains salient emotions (love, fear) as well as sensory states (gustation, pain). Awareness of said states, which generally includes a declarative component, may explain the insula's function in language processing. But then what of the insula's contribution to uncertainty, prediction errors and perception (Klein et al., 2013)? While error is generically viewed as a conflict or mistake, from a computational perspective it is simply a difference between two states. In a related manner, the insula constitutes a main component of the salience network in resting state fMRI (Menon and Uddin, 2010). Salience implies, at minimum, a deviation from a neutral state. We hypothesize that the insula's involvement in uncertainty reflects its role in mediating between upstream prediction errors (irrespective of origin) and declarative states. Thus, if an agent can label a state and the latter arises from a prediction error, we would expect insular involvement.

#### 4.7. Limitations

In this study we provide evidence for a functional role of the anterior insula with respect to uncertainty and surprise. We must note on some limitations the study presented however. First, anterior insula activation did not survive whole-brain family-wise error correction, only FDR, which presents a less stringent means to correct for multiple comparisons. Second, while efforts were made to create a novel perceptual task that matched a well-established financial task in its elicitation of uncertainty, several important differences remained. The financial uncertainty task presented a wider range of surprise values than the perceptual uncertainty task. Then, the perceptual uncertainty task included a potential surprise component unrelated to the stimulus value, in perceptual dominance times. A time-related surprise could be introduced in future studies by including a stimulus onset asynchrony between onsets of Card 1 and Card 2, to better approximate variability between surprising events. Finally, the perceptual uncertainty task demanded a motor response to a surprising event while the financial task did not. The motor response elicited anterior insular responses and may have compromised surprise-related insular activity.

#### 4.8. Inference and consciousness

Helmholtzian insights on inference hover over our principal question: does the brain infer reality in a general manner irrespective of stimulus type (visual, monetary) or end goal (adequate perception, or monetary gain)? The inferential process should be the cornerstone of our interaction with an uncertain environment, regardless of domain. Our aim was to exploit a parsimonious economic model of uncertainty in perceptual and financial decision-making to probe the inferential process assumed in both. Our perceptual task eschews key aspects of economic decisions: it is unconscious, involuntary and quasi-immediate. Nonetheless, the key element of uncertainty allows us to quantify the inferential process and extract measures of risk, as is commonly done in economic contexts. Crucially, perceptual uncertainty does not remain a mere theoretical quantity: we edify our hypothesis with an insular BOLD response, mimicking the neural correlates of (economic) uncertainty processing in previous studies. In examining the insula's functional role, we hypothesize that it is uniquely responsible for *conscious inference*, the result of an inference that can be recognized, as with a perceptual switch or the outcome of a gamble. Our results overall suggest that uncertainty can be quantified with a common framework across functional domains and further implicates a shared neural system.

#### Declaration of competing interest

The authors declare no competing financial interests.

#### CRedit authorship contribution statement

**Leyla Loued-Khenissi:** Writing - original draft, Investigation, Formal analysis, Methodology, Software. **Adrien Pfeuffer:** Data curation, Software, Investigation. **Wolfgang Einhäuser:** Funding acquisition, Writing - review & editing, Conceptualization, Validation, Formal analysis. **Kerstin Preuschoff:** Conceptualization, Funding acquisition, Writing - review & editing.

#### Acknowledgments

This work was supported by the Swiss National Science Foundation (135687) and the German Research Foundation (DFG; grant no.: EI 852/3-1).

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2020.116549>.

#### References

- Ackermann, H., Riecker, A., 2010. The contribution (s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214 (5–6), 419–433.
- Ardila, A., Bernal, B., Rosselli, M., 2014. Participation of the insula in language revisited: a meta-analytic connectivity study. *J. Neurolinguistics* 29, 31–41.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *Neuroimage* 26 (3), 839–851.
- Balleine, B.W., Delgado, M.R., Hikosaka, O., 2007. The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27 (31), 8161–8165.
- Bamiou, D.E., Musiel, F.E., Luxon, L.M., et al., 2003. The insula (Island of Reil) and its role in auditory processing: literature review. *Brain Res. Rev.* 42 (2), 143–154.
- Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision making. *Proc. Natl. Acad. Sci.* 107 (50), 21767–21772.
- Binder, J.R., Liebenthal, E., Possing, E.T., Medler, D.A., Ward, B.D., 2004. Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295–301.
- Boring, E.G., 1930. A new ambiguous figure. *Am. J. Psychol.* 42, 444–445.
- Brainard, D.H., Vision, S., 1997. The psychophysics toolbox. *Spat. Vis.* 10, 433–436.
- Brascamp, J.W., van Ee, R., Pestman, W.R., van den Berg, A.V., 2005. Distributions of alternation rates in various forms of bistable perception. *J. Vis.* 5 (4), 287–298.
- Bullmore, E., Sporns, O., 2012. The economy of brain network organization. *Nat. Rev. Neurosci.* 13 (5), 336.
- Brevers, D., Bechara, A., Hermoye, L., Divano, L., Kornreich, C., Verbanck, P., Noël, X., 2015. Comfort for uncertainty in pathological gamblers: a fMRI study. *Behav. Brain Res.* 278, 262–270.
- Benjamini, Hochberg, 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300.
- Carlson, J.M., Foti, D., Mujica-Parodi, L.R., Harmon-Jones, E., Hajcak, G., 2011. Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage* 57 (4), 1608–1616.
- Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36 (3), 181–204.
- Clark, A., 2017. Busting out: predictive brains, embodied minds, and the puzzle of the evidentiary veil. *Noûs* 51 (4), 727–753.
- Corradi-Dell'Acqua, C., Hofstetter, C., Vuilleumier, P., 2011. Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *J. Neurosci.* 31 (49), 17996–18006.
- Craig, A.D., Craig, A.D., 2009. How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10 (1).
- Critchley, H.D., Wiens, S., Rotshtein, P., Öhman, A., Dolan, R.J., 2004. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7 (2), 189–195.
- Combrisson, E., Jerbi, K., 2015. Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J. Neurosci. Methods* 250, 126–136.
- d'Acremont, M., Bossaerts, P., 2008. Neurobiological studies of risk assessment: a comparison of expected utility and mean-variance approaches. *Cognit. Affect Behav. Neurosci.* 8 (4), 363–374.
- Daniel, R., Pollmann, S., 2014. A universal role of the ventral striatum in reward-based learning: evidence from human studies. *Neurobiol. Learn. Mem.* 114, 90–100.
- Dayan, P., 1998. A hierarchical model of binocular rivalry. *Neural Comput.* 10 (5), 1119–1135.
- Doya, K., Ishii, S., Pouget, A., Rao, R.P. (Eds.), 2007. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. MIT press.
- Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and surprise determine neural population responses in the ventral visual stream. *J. Neurosci.* 30, 16601–16608. <https://doi.org/10.1523/JNEUROSCI.2770-10.2010>. PMID: 21147999.
- Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. U.S.A.* 201602413.
- Faraji, M., Preuschoff, K., Gerstner, W., 2018. Balancing new against old information: the role of puzzlement surprise in learning. *Neural Comput.* 30 (1), 34–83.
- Frässle, S., Sommer, J., Jansen, A., Naber, M., Einhäuser, W., 2014. Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J. Neurosci.* 34 (5), 1738–1747.
- Freeman, W.T., 1994. The generic viewpoint assumption in a framework for visual perception. *Nature* 368, 542–545.
- Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S., 1995. Spatial registration and normalization of images. *Hum. Brain Mapp.* 3 (3), 165–189.
- Friston, K., Adams, R., Perrinet, L., Breakspear, M., 2012a. Perceptions as hypotheses: saccades as experiments. *Front. Psychol.* 3, 151.
- Friston, K.J., Stephan, K.E., 2007. Free-energy and the brain. *Synthese* 159 (3), 417–458.
- Friston, K., 2010. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11 (2), 127–138.
- Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. *Neuroimage* 25 (3), 661–667.
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G., 2017. Active inference: a process theory. *Neural Comput.* 29 (1), 1–49.
- Friston, K., Thornton, C., Clark, A., 2012b. Free-energy minimization and the dark-room problem. *Front. Psychol.* 3, 130.
- Gasquoine, P.G., 2014. Contributions of the insula to cognition and emotion. *Neuropsychol. Rev.* 24 (2), 77–87.
- Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9 (4), 416–429.

- Gregory, R.L., 1997. Knowledge in perception and illusion. *Philos. Trans. R. Soc. Biol. Sci.* 352 (1358), 1121–1127.
- Grinband, J., Hirsch, J., Ferrera, V.P., 2006. A neural representation of categorization uncertainty in the human brain. *Neuron* 49, 757–763.
- Haefner, R.M., Berkes, P., Fiser, J., 2016. Perceptual decision-making as probabilistic inference by neural sampling. *Neuron* 90 (3), 649–660.
- Han, H., Glenn, A.L., 2018. Evaluating methods of correcting for multiple comparisons implemented in SPM12 in social neuroscience fMRI studies: an example from moral psychology. *Soc. Neurosci.* 13 (3), 257–267.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28 (22), 5623–5630.
- Heekeren, H.R., Marrett, S., Ungerleider, L.G., 2008. The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9 (6), 467.
- Heekeren, H.R., Marrett, S., Ruff, D.A., Bandettini, P.A., Ungerleider, L.G., 2006. Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proc. Natl. Acad. Sci.* 103 (26), 10023–10028.
- Ho, T.C., Brown, S., Serences, J.T., 2009. Domain general mechanisms of perceptual decision making in human cortex. *J. Neurosci.* 29 (27), 8675–8687.
- Hohwy, J., Roepstorff, A., Friston, K., 2008. Predictive coding explains binocular rivalry: An epistemological review. *Cognition* 108 (3), 687–701.
- Howard, I.P., 1996. Alhazen's neglected discoveries of visual phenomena. *Perception* 25 (10), 1203–1217.
- Hugrass, L., Crewther, D., 2012. Willpower and conscious percept: volitional switching in binocular rivalry. *PLoS One* 7 (4), e35963.
- Jones, C.L., Ward, J., Critchley, H.D., 2010. The neuropsychological impact of insular cortex lesions. *J. Neurol. Neurosurg. Psychiatry* 81 (6), 611–618.
- Kahneman, D., Tversky, A., 1982. Variants of uncertainty. *Cognition* 11 (2), 143–157.
- Kepecs, A., Mainen, Z.F., 2012. A computational framework for the study of confidence in humans and animals. *Philos. Trans. R. Soc. Biol. Sci.* 367 (1594), 1322–1337.
- Kettinger, Á., Hill, C., Vidnyánszky, Z., Windischberger, C., Nagy, Z., 2016. Investigating the group-level impact of advanced dual-echo fMRI combinations. *Front. Neurosci.* 10, 571.
- Klein, T.A., Ullsperger, M., Danielmeier, C., 2013. Error awareness and the insula: links to neurological and psychiatric diseases. *Front. Hum. Neurosci.* 7, 14.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., 2007. What's new in Psychtoolbox-3. *Perception* 36 (14), 1.
- Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27 (12), 712–719.
- Kuhnen, C.M., Knutson, B., 2005. The neural basis of financial risk taking. *Neuron* 47 (5), 763–770.
- Kroll, Y., Levy, H., Markowitz, H.M., 1984. Mean-variance versus direct utility maximization. *J. Financ.* 39 (1), 47–61.
- Leopold, D.A., Maier, A., Wilke, M., Logothetis, N.K., 1998. 13 Binocular Rivalry and the Illusion of Monocular Vision.
- Long, G.M., Toppino, T.C., 2004. Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol. Bull.* 130 (5), 748.
- Markowitz, H.M., March 1952. Portfolio selection. *J. Financ.* 7 (1), 77–91.
- Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* 214 (5–6), 655–667.
- Mohr, P.N., Biele, G., Heekeren, H.R., 2010. Neural processing of risk. *J. Neurosci.* 30 (19), 6613–6619.
- Naber, M., Frässle, S., Einhäuser, W., 2011. Perceptual rivalry: reflexes reveal the gradual nature of visual awareness. *PLoS One* 6 (6), e20910.
- Necker, L.A., 1832. LXI. Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 1 (5), 329–337.
- O'Connell, R.G., Shadlen, M.N., Wong-Lin, K., Kelly, S.P., 2018. Bridging neural and computational viewpoints on perceptual decision-making. *Trends Neurosci.* 41 (11), 838–852. <https://doi.org/10.1016/j.tins.2018.06.005>. Epub 2018 Jul 12.
- Parkkonen, L., Andersson, J., Hämäläinen, M., Hari, R., 2008. Early visual brain areas reflect the percept of an ambiguous scene. *Proc. Natl. Acad. Sci.* 105 (51), 20500–20504.
- Parr, T., Friston, K.J., 2017. The active construction of the visual world. *Neuropsychologia* 104 (July), 92–101.
- Payzan-Lenestour, E., Dunne, S., Bossaerts, P., Doherty, J.P.O., O'Doherty, J.P., 2013. The neural representation of unexpected uncertainty during value-based decision making. *Neuron* 79 (1), 191–201.
- Pearson, J., Tadin, D., Blake, R., 2007. The effects of transcranial magnetic stimulation on visual rivalry. *J. Vis.* 7 (7), 2.
- Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10 (4), 437–442.
- Pellicano, E., Burr, D., 2012. When the world becomes “too real”: a Bayesian explanation of autistic perception. *Trends Cogn. Sci.* 16 (10), 504–510.
- Peyron, R., Laurent, B., Garcia-Larrea, L., 2000. Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiol. Clin./Clin. Neurophysiol.* 30 (5), 263–288.
- Platt, M.L., Huettel, S.a., 2008. Risky business: the neuroeconomics of decision making under uncertainty. *Nat. Neurosci.* 11 (4), 398–403.
- Polonsky, A., Blake, R., Braun, J., Heeger, D.J., 2000. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nat. Neurosci.* 3 (11), 1153.
- Poser, B.A., Norris, D.G., 2009. Investigating the benefits of multi-echo EPI for fMRI at 7 T. *Neuroimage* 45 (4), 1162–1172.
- Preusschoff, K., Bossaerts, P., Quartz, S.R., 2006. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51 (3), 381–390.
- Preusschoff, K., Quartz, S.R., Bossaerts, P., 2008a. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28 (11), 2745–2752.
- Preusschoff, K., Quartz, S., Bossaerts, P., 2008b. Markowitz in the brain? *Rev. Écon. Polit.* 118 (1), 75–95.
- Preusschoff, K., Hart, B.M., Einhäuser, W., 2011. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Front. Neurosci.* 5, 115.
- Quartz, S.R., 2009. Reason, emotion and decision-making: risk and reward computation with feeling. *Trends Cogn. Sci.* 13 (5), 209–215.
- Rao, R.P., 2010. Decision making under uncertainty: a neural model based on partially observable markov decision processes. *Front. Comput. Neurosci.* 4.
- Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2 (1), 79.
- Ratcliff, R., 1979. Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86, 446–461.
- Rubin, E., 1921. *Visuell Wahrgenommene Figuren*. Gyldendalske Boghandel, Copenhagen, Berlin, London.
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. *J. Neurosci.* 30 (40), 13525–13536.
- Schmack, K., Schnack, A., Priller, J., Sterzer, P., 2015. Perceptual instability in schizophrenia: probing predictive coding accounts of delusions with ambiguous stimuli. *Schizophr. Res.: Cognition* 2 (2), 72–77.
- Schultz, W., 2010. Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6 (1), 24.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275 (5306), 1593–1599.
- Schwartenbeck, P., FitzGerald, T.H., Mathys, C., Dolan, R., Kronbichler, M., Friston, K., 2015. Evidence for surprise minimization over value maximization in choice behavior. *Sci. Rep.* 5, 16575.
- Singer, T., Critchley, H.D., Preusschoff, K., 2009. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13 (8), 334–340.
- Sladky, R., Friston, K.J., Tröstl, J., Cunningham, R., Moser, E., Windischberger, C., 2011. Slice-timing effects and their correction in functional MRI. *Neuroimage* 58 (2), 588–594.
- Small, D.M., 2010. Taste representation in the human insula. *Brain Struct. Funct.* 214 (5–6), 551–561.
- Sterzer, P., Kleinschmidt, A., 2007. A neural basis for inference in perceptual ambiguity. *Proc. Natl. Acad. Sci.* 104 (1), 323–328.
- Sterzer, P., Kleinschmidt, A., 2010. Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Struct. Funct.* 214 (5–6), 611–622.
- Sterzer, P., Kleinschmidt, A., Rees, G., 2009. The neural bases of multistable perception. *Trends Cogn. Sci.* 13 (7), 310–318.
- Strange, B.A., Duggins, A., Penny, W., Dolan, R.J., Friston, K.J., 2005. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw.* 18 (3), 225–230.
- Summerfield, C., De Lange, F.P., 2014. Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15 (11), 745.
- Summerfield, C., Tsetsos, K., 2012. Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Front. Neurosci.* 6, 70.
- Sundareswara, R., Schrater, P.R., 2008. Perceptual multistability predicted by search model for Bayesian decisions. *J. Vis.* 8 (5), 12.
- Thielscher, A., Pessoa, L., 2007. Neural correlates of perceptual choice and decision making during fear-disgust discrimination. *J. Neurosci.* 27, 2908–2917.
- Tricomi, E., Balleine, B.W., O'Doherty, J.P., 2009. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29 (11), 2225–2232.
- Uddin, L.Q., Nomi, J.S., Hébert-Seropian, B., Ghaziri, J., Boucher, O., 2017. Structure and function of the human insula. *J. Clin. Neurophysiol.: Off. Publ. Am. Electroencephalographic Soc.* 34 (4), 300.
- van Ee, R., Van Dam, L.C.J., Brouwer, G.J., 2005. Voluntary control and the dynamics of perceptual bi-stability. *Vis. Res.* 45 (1), 41–55.
- Vilares, I., Howard, J.D., Fernandes, H.L., Gottfried, J.A., Kording, K.P., 2012. Differential representations of prior and likelihood uncertainty in the human brain. *Curr. Biol.* 22 (18), 1641–1648.
- Volz, K.G., Schubotz, R.L., von Cramon, D.Y., 2005. Variants of uncertainty in decision-making and their neural correlates. *Brain Res. Bull.* 67 (5), 403–412.
- Von Helmholtz, H., 1867. *Handbuch der physiologischen Optik*. [Physiological Optics]. Leipzig: L. Voss).
- Von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, p. 625.
- Wang, M., Arteaga, D., He, B.J., 2013. Brain mechanisms for simple perception and bistable perception. *Proc. Natl. Acad. Sci.* 110 (35), E3350–E3359.
- Weinhammer, V., Stuke, H., Hesselmann, G., Sterzer, P., Schmack, K., 2017. A predictive coding account of bistable perception—a model-based fMRI study. *PLoS Comput. Biol.* 13 (5), e1005536.
- Weller, J.A., Levin, I.P., Shiv, B., Bechara, A., 2009. The effects of insula damage on decision-making for risky gains and losses. *Soc. Neurosci.* 4 (4), 347–358.
- Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., Rizzolatti, G., 2003. Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron* 40 (3), 655–664.
- Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. *Neuroimage* 15 (1), 1–15.
- Xue, G., Lu, Z., Levin, I.P., Bechara, A., 2010. The impact of prior risk experiences on subsequent risky decision-making: the role of the insula. *Neuroimage* 50 (2), 709–716.
- Zhou, Y.H., Gao, J.B., White, K.D., Merk, I., Yao, K., 2004. Perceptual dominance time distributions in multistable visual perception. *Biol. Cybern.* 90 (4), 256–263.