bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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-NC-ND 4.0 International license.

Electrophysiological indices reflect switches between Bayesian and heuristic strategies in perceptual learning

Daniel Bennett*
School of Psychological Sciences, The University of Melbourne

Karen Sasmita
School of Psychological Sciences, The University of Melbourne

Carsten Murawski
Department of Finance, The University of Melbourne

Stefan Bode
School of Psychological Sciences, The University of Melbourne

Abstract

Given finite cognitive resources, agents should allocate these to maximise desirable outcomes while minimising cognitive effort. This trade-off has often been studied as a competition between Bayesian inference and 'fast-and-frugal' heuristic strategies. An important open question in this regard is whether utilisation of Bayesian inference is dependent upon motivational state, and how this is reflected in the brain. We recorded electroencephalography from 23 participants performing a perceptual learning task with both monetary and a non-monetary instructive feedback conditions. Using model-based cluster analysis, we found that only participants who switched between a Bayesian and a heuristic strategy showed worse performance for instructive than monetary feedback, whereas participants who consistently employed Bayesian inference showed equivalent performance in both feedback conditions. This pattern of behavioural results was mirrored by differences in neural encoding of feedback in two event-related potential components: the P3, and the late positive potential. These findings suggest that use of Bayesian inference in perceptual learning may depend on motivational state.

* daniel.bennett@princeton.edu

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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-NC-ND 4.0 International license.

Introduction

Humans possess finite cognitive resources. In judgment and decision making these resources should be allocated so as to optimise decisions about behaviourally relevant outcomes, while minimising expenditure of cognitive resources on irrelevant or inconsequential tasks (Pitz & Sachs, 1984; Simon, 1976). Cognitive resource constraints are thought to provide a principled explanation for the finding that in many tasks humans rely on simple heuristics rather than adopting superior but more computationally demanding task strategies (Goldstein & Gigerenzer, 2002; Tversky & Kahneman, 1973, 1974). By producing reasonably accurate choices while consuming relatively few cognitive resources, heuristics can be a valuable tool to overcome cognitive resource constraints (Conlisk, 1996; Gigerenzer & Goldstein, 1996). However, the cognitive and neurophysiological factors affecting the use of optimal strategies versus resource-cheap heuristics remain unclear. In particular, one open question is how the presence of reward and motivational state might affect strategy choice (Achtziger, Alós-Ferrer, Hügelschäfer, & Steinhauser, 2015; Charness & Levin, 2005). This question is of particular importance given ongoing debate in educational psychology and personnel economics regarding the efficacy of using rewards to incentivise performance (Hidi, 2016; Lazear, 2000).

In learning research, one way that the trade-off between optimal and heuristic strategies has been conceptualised is as a competition between Bayesian inference, which is statistically ideal but computationally demanding, and a win-stay lose-shift (WSLS) heuristic (Bennett, Murawski, & Bode, 2015; Charness & Levin, 2005; Steyvers, Lee, & Wagenmakers, 2009). Whereas Bayesian inference involves repeatedly revising a complete prior belief distribution, the WSLS heuristic assumes that participants simply select choice options which have previously led to reinforcement, or shift to a new option if a previous choice is not reinforced (Robbins, 1952). Although Bayesian models often fit learning behaviour well overall, their goodness-of-fit deteriorates sharply as the cognitive demands of Bayesian inference increase (Payzan-LeNestour & Bossaerts, 2011), or when a WSLS heuristic conflicts with Bayesian inference (Achtziger et al., 2015; Charness & Levin, 2005). Moreover, even in studies where Bayesian models fit group-level data well, a substantial proportion of participants nevertheless made choices better explained by a WSLS heuristic (Bennett et al., 2015; Steyvers et al., 2009). It remains to be determined how these findings can be reconciled within Bayesian theories of cognition. Such theories, which include both weak and strong variants, claim that much of human behaviour can be explained as a form of Bayesian inference (e.g., Knill & Pouget, 2004; Chater & Oaksford, 2008; Friston et al., 2015).

One possible explanation for these findings is that motivational factors may affect the use of Bayesian versus heuristic strategies. Given the greater cognitive demands of Bayesian inference compared with simple heuristics, participants may require more task motivation to employ a Bayesian strategy. As a result, individual differences in motivational state are likely to result in the utilisation of different task strategies and the recruitment of different neural processes by different participants. One corollary of this explanation is that, since feedback delivered in the form of monetary reward is thought to enhance participants' cognitive control (Fröber & Dreisbach, 2014; Jimura, Locke, & Braver, 2010), providing rewarding performance feedback should incentivise the use of Bayesian inference strategies

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

in complex task environments.

In the present study, we investigated the effect of motivational state on the use of Bayesian versus heuristic strategies in a perceptual learning task with graded feedback. In this task, feedback was delivered in the form of either monetary reward or instructional directives. Importantly, feedback values were constrained such that the exact amount of information provided by feedback was identical across the two feedback conditions. Then, to assess the question of strategy selection in a principled way, we formulated competing computational cognitive models implementing Bayesian, heuristic, or mixed strategies, and compared the predictions of these models with behavioural data. In order to elucidate the neural mechanisms underlying selection of Bayesian versus heuristic strategies, we investigated the effect of feedback condition on three event-related potential (ERP) components associated with learning and/or reward processing: the P3 (Polich, 2007), the feedback-related negativity (FRN; Yeung & Sanfey, 2004), and the late positive potential (LPP; Ito, Larsen, Smith, & Cacioppo, 1998). Finally, using a model-based cluster analysis, we investigated how neural encoding of feedback differed between participant subgroups employing different task strategies.

Method

Participants

Twenty-three participants were recruited from among students of the University of Melbourne, Australia (mean age = 23.40; age range 19-31; 17 female, six male). Participants were right-handed and had normal or corrected-to-normal visual acuity. Exclusion criterion was a medical history of any neurological disorder. Informed consent was acquired from all participants in accordance with the Declaration of Helsinki, and approval was obtained from the University of Melbourne Human Research Ethics Committee (ID 1339694). Participants received monetary compensation for participation (mean = $25.24; SD = 4.05) that was proportional to task winnings in the monetary feedback condition only (see below). For all participants, total remuneration value was within the range AUD $20-30.

Four participants were excluded from analysis of EEG data: one because of an excessive number of artefacts (more than 80 percent of trials affected), one because of a failure of the eyeblink artefact removal routine, and two additional participants because of computer error during EEG acquisition. Final EEG analyses were therefore performed on data provided by 19 participants (mean age = 23.75; age range 19-31; 13 female, six male).

Behavioural task

While EEG was recorded, participants performed a perceptual learning task modified from a previous study by Bennett et al. (2015). This task required participants to use visually presented feedback in sequential trials to learn the target contrast of a greyscale checkerboard stimulus (see Figure 1A). The present study employed a novel variant of this task in which feedback regarding the target contrast could be either monetary (as in the original paradigm) or instructive.

In each trial, the checkerboard was presented for up to 30 seconds (see Figure 1A), during which time its contrast changed linearly (alternately increasing and decreasing, changing direction at upper/lower contrast bounds). Initial contrast, initial direction of contrast

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

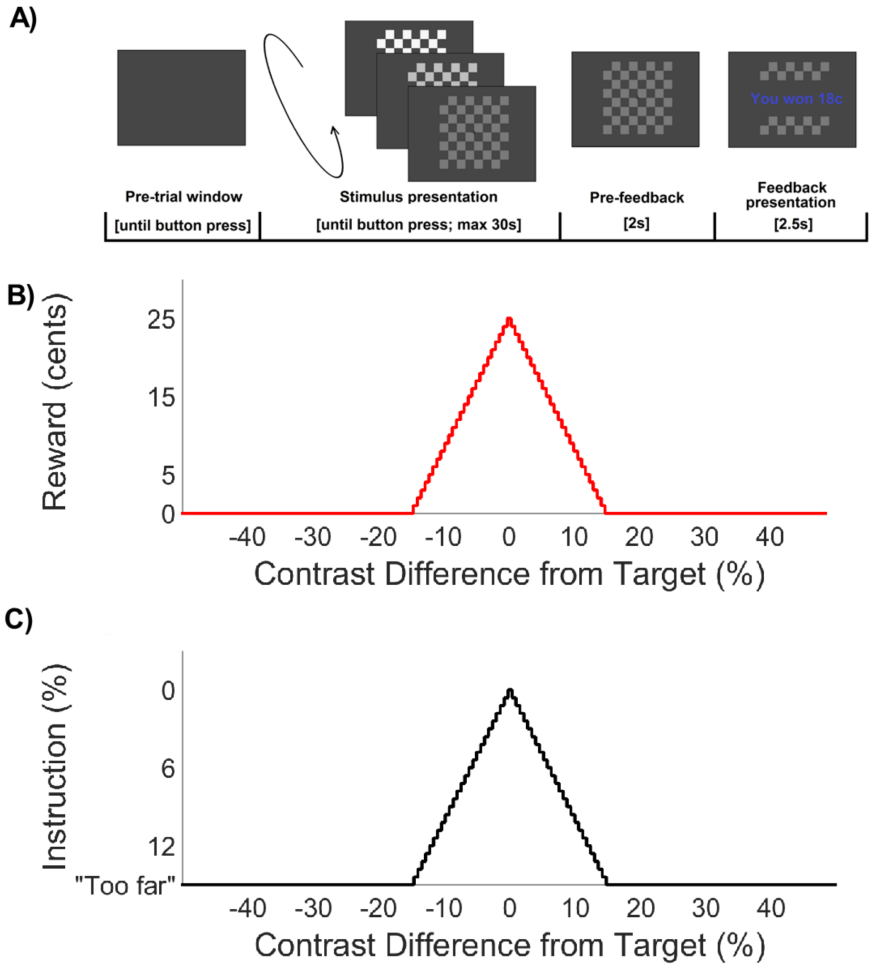

Figure 1. (A) Trial schematic. Following a self-paced button press, a checkerboard stimulus was presented with a linearly changing contrast. The participant could at any time select the contrast displayed on screen by pressing a button with the right index finger. The trial continued until a button was pressed, or until stimulus duration exceeded 30 seconds. Following the participant's choice, the selected contrast remained on screen for two seconds, after which time the monetary or instructive feedback associated with the chosen contrast was displayed for 2.5 seconds. In the event that no button was pressed within 30 seconds, feedback was a message reminding the participant of the task instructions. (B) Feedback mapping for monetary feedback condition. The mapping was a symmetrical triangular function with a centre of zero percent contrast difference, a half-width of 15 percent contrast difference, and a height of 25 cents. As such, received reward was maximal when the participant responded at the target contrast, and decreased with increasing difference of chosen contrast from the target. Reward was zero for responses at greater than 15 percent distance. Feedback received was rounded to the nearest whole-cent value. (C) Feedback mapping for the instructive feedback condition. For responses at less than 15 percent difference from the target, participants were informed of the difference between the chosen contrast and the target (rounded to the nearest of 49 equally spaced values, in order to match precisely the step size of the monetary condition's feedback mapping). For responses at greater than 15 percent difference from the target, participants were informed only that their response was "too far" from the target (equivalent to the monetary condition's zero cent feedback).

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

change (increasing/decreasing), and rate of change were randomised on each trial using the same parameters as in Bennett et al. (2015). At any time during stimulus presentation, the participant could choose the contrast displayed on screen by pressing a button with the right index finger. After a delay in which the chosen contrast remained on screen, participants received feedback regarding their chosen contrast. In the monetary condition, this feedback was presented in the form of monetary reward (e.g. "You won 15 cents") according to a triangular function M of the distance between the chosen and the target contrast (see Figure 1B). Responses closer to the target contrast earned proportionally more (up to a maximum of 25 cents per trial, rounded to the nearest integer), and participants received zero reward for responses at greater than 15 percent distance from the target:

$$M(r_t, x_t) = \begin{cases} \left\| 25 - \frac{5|r_t - x_t|}{3} \right\|, & |r_t - x_t| \leq 15 \\ 0, & |r_t - x_t| > 15 \end{cases} \tag{1}$$

where t is the trial number, r_t is the target contrast on trial t, and x_t is the participant's chosen contrast on trial t. Double bars denote rounding to the nearest integer.

By contrast, feedback in the instructive condition took the form of an explicit instructional directive informing the participant of the distance between their chosen contrast and the target (e.g. "You were 11.25% away from the target"; see Figure 1C). For responses at greater than 15 percent distance from the target, participants were informed only that their response was 'too far' from the target. As such, in the instructive feedback condition the function reward mapping function M from Equation 1 was replaced with the instruction mapping function I:

$$I(r_t, x_t) = \begin{cases} |r_t - x_t|, & |r_t - x_t| \leq 15 \\ \text{"Too far"}, & |r_t - x_t| > 15 \end{cases} \tag{2}$$

Crucially, in order to ensure strict equivalence in feedback information between instructive and monetary feedback, instructive feedback values were constrained to follow an equivalent functional form to monetary feedback (compare Figures 1B and 1C). This was done by rounding instructive feedback values to the nearest value in the set $\{0, 0.625, 1.25, 1.875, ...15\}$. For any given sequence of choices, therefore, feedback in the two conditions provided identical information regarding the target contrast. Consequently, any differences in task performance between instructive and monetary conditions cannot be attributed to differences in the information content of feedback.

Prior to the task, participants were trained in interpretation of feedback in both feedback conditions, and testing commenced only when satisfactory levels of task understanding were displayed. Participants then completed 14 blocks of the task in total over approximately 50 minutes. Each block had a different target contrast, selected pseudo-randomly from the interval $[25\%, 85\%]$. Monetary and instructive conditions were presented in seven consecutive blocks each, with condition order counterbalanced across participants. Each block continued until cumulative checkerboard presentation duration for the block exceeded three minutes, or until 25 trials were completed, whichever occurred sooner. As a result, the number of trials per block varied, ensuring that participants could not rush through the task in an attempt to trade off experiment duration against monetary winnings. Upon receiving feedback, participants were not informed of the exact numerical contrast level of their

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

choice; instead, the checkerboard remained on-screen at the chosen contrast while feedback was presented. As a result, learning was necessarily affected by perceptual uncertainty regarding the identity of the chosen contrast.

Stimuli were presented using a Sony Trinitron G420 CRT monitor at a framerate of 120 Hz. During task performance, participants were seated comfortably in a darkened room, using a chin rest at a distance of 77 cm from the screen. Checkerboard stimuli were 560 $\times$ 560 pixels in size, measuring 19.5 $\times$ 19.5 cm on the screen and subtending a visual angle of 14.43 $\times$ 14.43°. Responses were recorded using a five-button Cedrus Response Box. All other task parameters were identical to those employed by Bennett et al. (2015), with the exception that the checkerboard in the present task did not phase-reverse, and therefore had a smoothly changing (rather than flickering) appearance.

Computational models of behaviour

We tested four competing computational cognitive models of the task by generating all possible configurations of Bayesian and heuristic strategies according to feedback condition. This model configuration permitted us to test formally several possible ways in which participants might have switched, or not have switched, between Bayesian and heuristic strategies as a function of feedback condition (see below for formal specifications and choice rules for the Bayesian and heuristic strategies).

The first model, termed the 'Consistent Bayesian' (CB) model, assumed that participants employed a Bayesian inference strategy irrespective of the feedback condition. Similarly, the second model, the 'Consistent Heuristic' (CH) model, assumed that participants employed a WSLS heuristic strategy in both the monetary and the instructive feedback condition. A third possibility was that participants might employ a Bayesian inference strategy when monetary feedback was provided, but switch to a WSLS heuristic strategy in blocks with instructive feedback. Such a model implies the use of a more computationally demanding strategy in the presence of reward and a simpler heuristic strategy in the absence of reward, and was therefore termed the 'Incentive-Compatible Switching' (ICS) model. Finally, for the fourth model we also considered the possibility that participants might employ a WSLS heuristic strategy in monetary feedback blocks, and a Bayesian inference strategy in instructive feedback blocks. This model was termed the 'Incentive-Incompatible Switching' (IIS) model, and can be thought of as corresponding to the idea of "choking under pressure" (see also Achtziger et al., 2015; Baumeister, 1984), in which the presence of monetary incentives produces a decrement in performance.

Each of these four models was parameterised by the perceptual uncertainty parameter σ (see below); this parameter was permitted to vary between participants, but was fixed to take the same value across different feedback conditions for each participant, since neither perceptual stimuli nor observation environment varied across feedback conditions. Constraining the σ parameter across feedback conditions was done for identification purposes, to ensure that differences in task performance between feedback conditions could not simply be accounted for by changes in a perceptual uncertainty parameter. One prominent criticism of Bayesian models in psychology and neuroscience is that flexible parameterisations of Bayesian models permit qualitatively and quantitatively distinct patterns of behaviour to be described within an identical model architecture (Bowers & Davis, 2012). The constraints on σ applied in the present study ensured that the goodness-of-fit of each of the four models

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

described below gave an unbiased estimate of the likelihood of each behavioural strategy. For each model, σ was estimated using maximum likelihood estimation as implemented in the MATLAB Optimization Toolbox (The Mathworks, Natick, MA) and choice likelihood functions as specified by Bennett et al. (2015). Model fitting was performed with multiple different initial parameter values for each participant to ensure that identified values of sigma corresponded to global rather than local optima.

The relative performance of these four models also informed subsequent ERP analyses. By comparing model fits for individual participants we derived an index of which model strategy provided the best explanation of each participant's data. Groups of participants who all used the same strategy were then clustered into subgroups for subsequent ERP analyses. Differences in neural encoding of feedback between subgroups were, therefore, interpretable with respect to the different behavioural strategies employed by different participants. This model-based clustering analysis is a principled alternative to more traditional model-free cluster analysis algorithms, which have been criticised for clustering data in order to maximise intra-cluster homogeneity in a way that may not result in psychologically meaningful differences between participant subgroups (Fraley & Raftery, 1998; Meehl, 1992). By contrast, a model-based clustering procedure allows for principled segregation of participants into subgroups representing different computational models of task performance. As such, different subgroups necessarily correspond to meaningful and distinct psychological constructs. In the present study, this approach meant that differences between subgroups in neural encoding of feedback could be readily interpreted with reference to the employment of different behavioural strategies. This approach can be considered a classification-based counterpart of Bayesian model selection (Marković, Gläscher, Bossaerts, O'Doherty, & Kiebel, 2015; Stephan, Penny, Daunizeau, Moran, & Friston, 2009), which also takes into account within-participants variability in the likelihood of different computational models.

It should also be noted that the model of Bayesian inference implemented in the present study is only one among many models capable of approximating Bayesian updating for feedback-based learning under uncertainty (e.g., Behrens, Woolrich, Walton, & Rushworth, 2007; Mathys, Daunizeau, Friston, & Stephan, 2011; Nassar, Wilson, Heasly, & Gold, 2010). The statistical mechanics of Bayesian updating are not unique to the model used in the present study; as such, this model should be considered as representing a specific implementation of broad computational principles which are applicable to a large number of distinct model architectures.

Bayesian strategy. To model choices under the Bayesian inference strategy, we estimated beliefs using a Bayesian grid estimator (Moravec, 1988) as described and implemented for the perceptual learning task used in the present study by Bennett et al. (2015). This estimator calculated a probabilistic estimate of participants' beliefs regarding the level of the target contrast in each trial, and used this belief distribution to estimate choice likelihoods. Formally, beliefs were described by a probability mass function θ over a contrast space discretised into J equally sized bins, where the value of the function θ at each bin represented the participant's subjective probability that the target contrast (denoted r_t) fell within bin j on trial t. On each trial t, participants observed the feedback ft after the choice of contrast bin x_t, determined according to the monetary and instructive feedback mapping functions M and I as specified by Equations 1 and 2, respectively. Belief estimates were initialised in

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

each block as a discrete uniform distribution, representing participants' a priori uncertainty regarding the target contrast level. This belief distribution was then updated sequentially according to Bayes' Rule as feedback was received, such that the posterior distribution of trial t formed the prior distribution for trial $t+1$:

$$\theta_{t+1}(j) = \frac{\theta_t(j) Pr(f_t, x_t | r_t \in j)}{Pr(f_t, x_t)} \tag{3}$$

The left-hand side of Equation 3 represents the posterior belief distribution for contrast bin j following trial t, and is calculated by multiplying the participant's prior belief that the target contrast fell within bin j, $\theta_t(j)$ by the likelihood of observing the choice/feedback pair (f_t, x_t) if the target were in bin j, $Pr(f_t, x_t | r_t \in j)$, and dividing by the marginal likelihood of the update, $Pr(f_t, x_t)$.

As described above, variability in task performance between participants was captured by the perceptual uncertainty parameter σ. Formally, σ represents the standard deviation of the Gaussian noise affecting belief updates after feedback receipt, such that larger values of σ indicate a greater degree of noise in the updating process, and therefore more imprecise belief updates. Since participants were not informed of exactly what contrast they had chosen, but had to estimate this chosen contrast from the visual display, this perceptual uncertainty therefore also results in a Gaussian prior over chosen contrast. For a complete discussion of the mathematical role of σ in the Bayesian updating model see Bennett et al. (2015).

To estimate choice likelihood, this model used a probability of maximum utility choice rule (cf. Speekenbrink & Konstantinidis, 2015), whereby contrast bins with a higher probability of containing the target contrast had a proportionally higher probability of being chosen, subject to response uncertainty during choice:

$$Pr(x_t) = \frac{\theta * G_0 [J]}{k} \tag{4}$$

As such, on each trial the choice likelihood probability mass function was determined by convolving the prior belief distribution θ by the uncertainty function G_0 over the set of contrast bins J, where k is a normalisation constant and square parentheses denote the domain of convolution. Intuitively, this response model implies that response probabilities are derived by the addition of Gaussian noise to the target contrast distribution J. The uncertainty function G_0 was a zero-mean Gaussian function of the contrast difference between the true chosen contrast x_t and each bin x_j of the distribution θ. This function was also parameterised by σ (truncated to the available range of contrasts):

$$G_0(x_t, x_k, \sigma) \sim \mathcal{N}(x_t - x_j, \sigma) \tag{5}$$

WSLS Heuristic Strategy. In contrast to the Bayesian inference strategy, the WSLS heuristic strategy did not assume that participants attempted to infer the location of the target contrast. Instead, behaviour under the WSLS heuristic was assumed to be driven by a one-trial memory, such that participants' behaviour on a given trial was a function of whether or not they had received reinforcement on the preceding trial (Robbins, 1952). Specifically, the model assumed that participants attempted to repeat the choice made on the previous trial if this choice had resulted in any level of reinforcement (defined as

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

any monetary reward amount in the monetary feedback condition, and as any numerical instruction amount in the instructive feedback condition). The WSLS model assumed that participants shifted randomly to a new contrast if they received no reinforcement on trial $t-1$, or at the start of a new block. This gives the following choice probability function:

$$Pr(x_t) = \begin{cases} \frac{(\delta(j-x_{t-1})*G_0)[J]}{k}, & f_{t-1} > 0 \\ \\ \frac{1}{J}, & \text{otherwise} \end{cases} \tag{6}$$

where G_0 is defined as per Equation 5, J is the number of bins in the belief distribution, δ is the Dirac delta function, and k is a normalisation constant. See Bennett et al. (2015) for a full formal characterisation of the WSLS heuristic as implemented for the perceptual learning task used in the present study.

Note that perceptual uncertainty in the WSLS choice rule was implemented in an identical manner to the Bayesian inference choice rule, using the same zero-mean Gaussian function. This allowed for an identical specification of perceptual uncertainty across all four models, thereby ensuring that predicted choice probabilities were directly comparable between models.

EEG data acquisition

The electroencephalogram was recorded from 35 Ag/AgCl active scalp electrodes (Fp2, AF7, AF3, AFz, AF4, AF8, F5, F1, Fz, F2, F4, F6, FC1, FCz, FC4, FC6, C5, C3, Cz, C4, CP5, CP3, CP1, CPz, CP6, P5, P1, Pz, P4, P6, POz, PO8, O1, Oz, Iz in the International 10-20 System). Electrodes interfaced with a BioSemi ActiveTwo 64-channel system running ActiView acquisition software, and used an implicit reference during recording. Due to technical problems with electrode hardware, not all 64 channels could be recorded for all participants. Therefore, based on previous (Bennett et al., 2015) and planned analyses, data was acquired from prespecified channels of interest, including all fronto-central and centro-parietal midline electrodes. All electrode channels included in subsequent event-related potential (ERP) analyses were recorded without issue for all participants, and data quality was not compromised. Data were linearly detrended and re-referenced offline to an average of left and right mastoid electrodes. The vertical and horizontal electrooculogram (EOG) were recorded from electrodes infraorbital and horizontally adjacent to the left eye. EEG was recorded at a sampling rate of 512 Hz.

Preprocessing of data was performed using a semi-automated preprocessing pipeline (cf. Bode, Bennett, Stahl, & Murawski, 2014; Brydevall, Bennett, Murawski, & Bode, 2017). Data were first manually screened to exclude epochs contaminated by skin potential or muscle artefacts. Using a linear FIR filter, data were then highpass filtered at 0.1 Hz, lowpass filtered at 70 Hz, and notch filtered at 50 Hz to remove background electrical noise. Epochs were generated consisting of data from 1500 milliseconds before to 1500 milliseconds after feedback presentation. An independent components analysis (ICA) as implemented in the EEGLAB toolbox (Delorme & Makeig, 2004) was performed on the resulting dataset to identify and remove components related to eye movements and eye-blink artefacts. Finally, an automatic artefact screening procedure excluded all epochs from analysis in which maximum/minimum amplitudes exceeded $\pm 200\ \mu$V.

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

ERP data analysis

We assessed three ERP components: the P3, the feedback-related negativity (FRN), and the late positive potential (LPP). Component amplitudes were calculated using estimation routines implemented in the ERPlab plugin (Lopez-Calderon & Luck, 2014), time-locked to feedback presentation on each trial and baseline-corrected from 0 to 500 milliseconds pre-feedback.

P3 amplitude was calculated as the largest positive peak in the window from 250-550ms post-feedback at the frontocentral and centroparietal midline electrodes AFz, Fz, FCz, Cz, and CPz (Bennett et al., 2015). This time window allowed us to estimate peak amplitude within a symmetrical window about the peak of the P3 as identified in grand average waveforms. Past research suggests that P3 amplitude is an index of individuals' revision of probabilistic beliefs (Kolossa, Kopp, & Fingscheidt, 2015), and we therefore investigated whether P3 amplitude varied between Bayesian belief updating and heuristic strategies.

At the same midline electrodes as the P3 analysis, FRN amplitude was calculated as the peak-to-peak distance between the most negative peak in the window from 200 to 550 milliseconds and the immediately preceding positive peak (Achtziger et al., 2015; Frank, Woroch, & Curran, 2005; Yeung & Sanfey, 2004). A peak-to-peak measure of the FRN was used rather than a mean amplitude measure to ensure that estimates of FRN and P3 amplitude were statistically independent of one another. FRN amplitude was investigated because of its importance as an index of outcome evaluation in reinforcement learning and feedback processing (Frank et al., 2005; Yeung & Sanfey, 2004; Holroyd & Coles, 2002).

Finally, LPP amplitude was calculated as the mean voltage within the window from 550 to 900 milliseconds post-feedback at the centro-parietal midline electrodes Cz, CPz, and Pz (Hajcak, Dunning, & Foti, 2009; Ito et al., 1998). This time window was chosen both to accord with previous literature (e.g., Keil et al., 2002), and to ensure that P3 and LPP analysis windows did not overlap. In research studying the processing of emotional stimuli, LPP amplitude is thought to differentially encode positive and neutrally valenced stimuli (Keil et al., 2002; Schupp et al., 2000); we therefore sought to investigate whether LPP amplitude differed between monetary and instructive feedback conditions.

ERP analyses investigated the neural correlates of differential processing of monetary and instructive feedback. Where formal comparison of computational cognitive models indicated the presence of participant subgroups using distinct performance strategies, we investigated interactions between model-derived participant subgroups and feedback condition. This allowed us to identify electrophysiological indices associated with the use of different behavioural strategies in different subgroups.

Results

Behavioural results

Participants completed a variable number of trials per block (mean = 17.57; SD = 2.70). A paired-samples t-test indicated that the average number of trials completed per block did not differ between instructive and monetary conditions ($t(22) = 0.33$, $p = .74$).

Behavioural performance was quantified by choice error, defined as the absolute contrast difference between the chosen contrast and the target contrast on each trial. We investigated differences in choice error as a function of trial number and feedback condition using linear

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

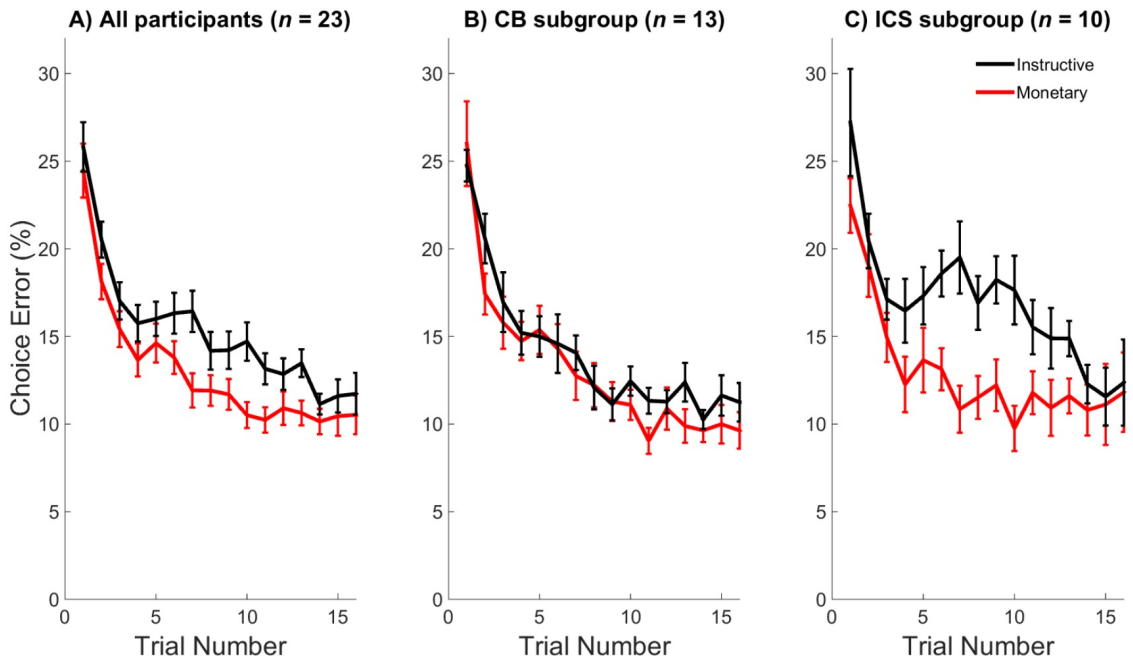

Figure 2. (A) Overall task performance among all participants ($n = 23$). All participants' mean choice error (measured as absolute difference between chosen and target contrasts) as a function of feedback condition (red: monetary feedback; black: instructive feedback) and trial number. Only trials 1-15 are shown, since data was not available for all participants for trials 16 – 25. Error bars represent the standard error of the mean. (B) Overall behavioural results among the subgroup of participants whose choices were best fit by the Consistent Bayesian (CB) model in a two-model comparison ($n = 13$). (C) Overall behavioural results among the subgroup of participants whose choices were best fit by the Incentive-Compatible Switching (ICS) model in a two-model comparison ($n = 10$).

mixed-effects analysis with feedback condition and trial number as fixed effects. Results indicated a significant main effect of trial number ($F(24,60.95) = 13.09$, $p = 4 \times 10^{-16}$), with performance improving over time within each block (see Figure 2A), and a significant main effect of feedback condition ($F(1,9.29) = 20.07$, $p = .001$), driven by better overall performance in the monetary than the instructive feedback condition. In addition, we observed a significant interaction between feedback condition and trial number ($F(24, 60.95) = 2.89$, $p = .0004$). This effect is likely to have been driven by greater differences between monetary and instructive feedback conditions in mid- and late-block trials, rather than in block-initial trials. Such a pattern stands to reason, since participants began each block with no a priori knowledge regarding the target contrast, and were as likely to make a correct as an incorrect initial guess regardless of feedback condition.

Computational model results

Using standard model comparison techniques, we next determined which of the computational models defined above provided the best account of choices across participants. Table 1 presents Bayesian Information Criterion (BIC) values for each of the four mod-

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

els. Results showed that, as in a previous study using this perceptual learning paradigm, the CB model, which assumed that participants adopted a Bayesian inference strategy in both feedback conditions, provided the best fit to data across all participants (Bennett et al., 2015). However, further examination of model fits for individual participants using participant-specific BIC values (see 'n best fit' columns) revealed that there was considerable variability in the best-fitting model across participants. Indeed, in spite of providing the best overall fit to data, the CB model was the best-fitting behavioural model for fewer than half (n = 11) of all participants considered separately. This strongly suggests the presence of inter-individual heterogeneity in task strategy. Closer inspection of individual model fits revealed that the second-best-fitting model overall, the ICS model, was the best-fitting model for approximately an additional third of participants. This model assumed that participants switched between a Bayesian inference strategy in the monetary feedback condition and a WSLS heuristic strategy in the instructive feedback condition.

Table 1
Summary of computational model fits to 5651 choices by 23 participants.

Model	n parameters	-LL[1]	BIC[2]	n best fit (4 models)	n best fit (2 models)
CB[3]	1	26094.18	52387.07	11	13
ICS[4]	1	26178.47	52555.66	7	10
IIS[5]	1	26254.21	52707.13	2	-
CH[6]	1	26344.96	52888.63	3	-

[1] Negative log likelihood [2] Bayesian Information Criterion [3] Consistent Bayesian model
[4] Incentive-Compatible Switching model [5] Incentive-Incompatible Switching model
[6] Consistent Heuristic model

As such, more than three-quarters of participants were best-fit by either the CB or the ICS model. Furthermore, for four of the five participants best fit by one of the other two models (IIS or CH), the second-best-fitting model was either the CB or the ICS model, such that the CB and ICS models were together either the first- or second-choice model for 22 of the 23 participants. Given the marked superiority of the CB and ICS models, therefore, we divided participants into two subgroups using a two-model comparison of the CB and ICS models (Fraley & Raftery, 1998). This model-based cluster analysis produced a two-model solution in which the principle of parsimony was balanced against the evident heterogeneity in individual task strategies detailed in Table 1.

We termed the two approximately evenly sized participant subgroups resulting from this two-model comparison the 'CB' (n = 13) and 'ICS' (n = 10) subgroups respectively. Subgroup membership was included as a between-subjects grouping variable in all subsequent behavioural and ERP analyses, in order to investigate whether patterns of learning and neural responses to feedback differed between participant subgroups. Furthermore, since the order in which feedback conditions were presented was counterbalanced across participants, we performed an additional control analysis to ensure that the two behavioural subgroups identified by model comparison were not merely a reflection of between-participants differences in condition order. A chi-square test of independence revealed no relationship

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

between condition order and model subgroup ($\chi^2(1) = 0.73$, $p = .39$); as such, there was no evidence to suggest that the classification of participants into behavioural subgroups was related to the order in which participants completed the instructive and monetary feedback conditions.

Next, we sought to determine whether behavioural strategy (as indexed by subgroup membership) was associated with different levels of overall task performance. We used a 2 × 2 repeated-measures ANOVA with within-groups factor of condition (instructive, monetary) and between-groups factor of model subgroup (CB, ICS), and mean choice accuracy across all trials as a dependent variable. We found a significant effect of feedback condition on accuracy ($F(1,21) = 20.88$, $p = .00017$, $\eta_p^2 = .50$), indicating better overall performance in the monetary condition than the instructive condition. Crucially, we also found a significant interaction between reward condition and model subgroup ($F(1,21) = 8.13$, $p = .01$, $\eta_p^2 = .28$). Follow-up paired-samples t-tests revealed that this interaction was driven by a significantly better overall performance in the monetary than the instructive condition in the ICS subgroup ($t(9) = -4.01$, $p = .003$; see Figure 3C), but not in the CB subgroup, ($t(12) = -1.66$, $p = .12$; see Figure 3B). In addition, we observed a non-significant trend toward better performance overall among the CB subgroup than among the ICS subgroup ($F(1,21) = 3.71$, $p = .07$).

ERP results

We next investigated whether any of the three identified ERP components displayed an interaction between feedback condition and behavioural subgroup that might account for the analogous interaction observed in behavioural data (see Figure 3). This analysis allowed us to identify electrophysiological indices which were associated with the differential relative performance between monetary and instructive feedback in participants who consistently adopted a Bayesian inference strategy (the CB model subgroup), compared with participants who switched between a Bayesian inference strategy and a heuristic strategy in different feedback conditions (the ICS subgroup). Scalp maps for the P3 and LPP analysis windows are presented in Figure 4.

P3. A 5×2×2 repeated-measures ANOVA with within-groups factors of electrode (AFz, Fz, FCz, Cz, CPz) and feedback condition (instructive, monetary) and between-groups factor of model subgroup (CB, ICS) revealed a significant main effect of electrode on P3 amplitude ($F(1.50, 25.56) = 13.94$, $p = .0002$, $\eta_p^2 = .45$), as well as a significant interaction between feedback condition and model subgroup on P3 amplitude ($F(1,17) = 4.96$, $p = .04$, $\eta_p^2 = 0.23$; see Figure 5A). Follow-up paired-samples t-tests marginalised across electrodes revealed that this interaction was driven by a significantly larger P3 amplitudes for monetary than instructive feedback in the ICS subgroup ($t(8) = -2.88$, $p = .02$), but not in the CB subgroup ($t(9) = .29$, $p = .78$). This indicates that P3 amplitudes differed between feedback conditions solely for participants who switched between a Bayesian strategy in the monetary feedback condition and a heuristic strategy in the instructive feedback condition; by contrast, there was no difference in P3 amplitudes for participants who employed a Bayesian strategy in both feedback conditions.

Finally, we also observed a non-significant trend toward a main effect of feedback condition on P3 amplitude ($F(1,17) = 3.32$, $p = .09$). No other main effects or interactions were statistically significant, all $p > .10$.

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

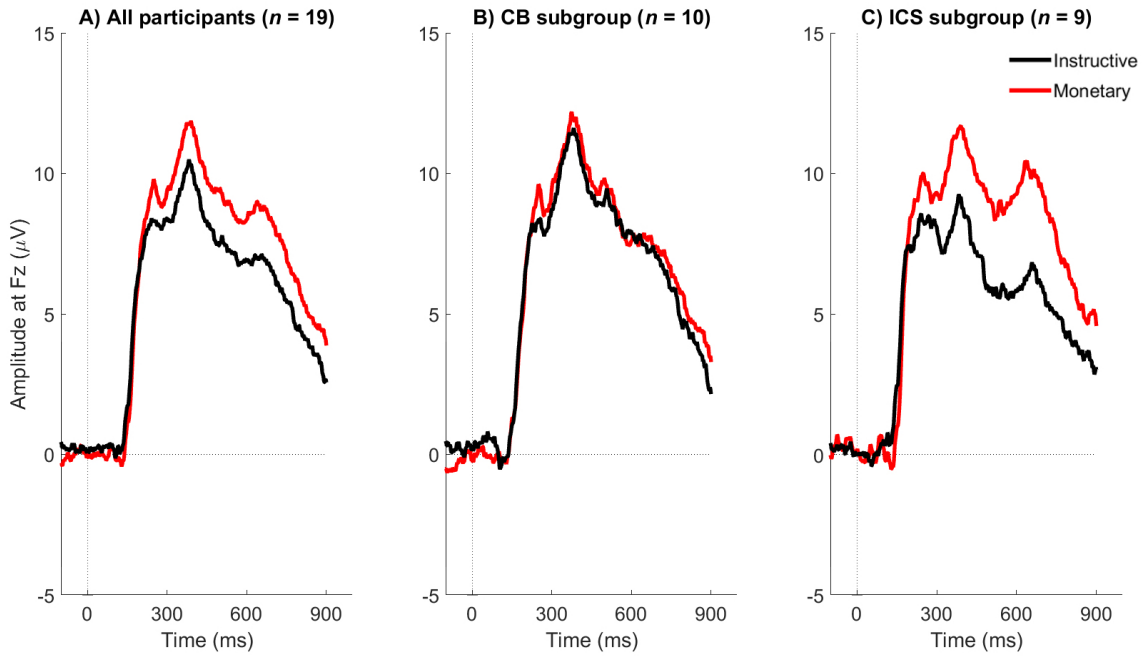

Figure 3. Grand average feedback-locked ERP waveforms at electrode Fz, grouped by feedback condition (red: monetary feedback; black: instructive feedback). Time 0 denotes the presentation of performance feedback. (A) Grand average waveforms across all participants. (B) Grand average waveforms across the subgroup of participants whose choices were best fit by the Consistent Bayesian (CB) model in a two-model comparison ($n = 10$). (C) Grand average waveforms across the subgroup of participants whose choices were best fit by the Incentive-Compatible Switching (ICS) model in a two-model comparison ($n = 9$).

FRN. A 5×2×2 repeated-measures ANOVA with within-groups factors of electrode (AFz, Fz, FCz, Cz, CPz), condition (instructive, monetary) and between-groups factor of model subgroup (CB, ICS) revealed a significant main effect of feedback condition on FRN amplitude ($F(1,17) = 5.92$, $p = .03$, $\eta_p^2 = .26$), with larger FRNs elicited by instructive than rewarding feedback. There was a non-significant trend toward an effect of model subgroup on FRN amplitude ($F(1,17) = 4.20$, $p = 0.06$), with numerically larger FRN amplitudes for the ICS than the CB subgroup (see Figure 5B). There was no interaction between feedback condition and model subgroup, and no other main effects or interactions were significant, all $p > .10$.

LPP. A 3×2×2 repeated measures ANOVA with within-groups factors of electrode (CPz, Pz, Cz) and feedback condition (instructive, monetary) and between-groups factor of model subgroup (CB, ICS) revealed a significant main effect of feedback condition on LPP amplitude ($F(1,17) = 21.38$, $p = .0002$, $\eta_p^2 = .56$), as well as a significant interaction between feedback condition and model subgroup ($F(1,17) = 4.85$, $p = .04$, $\eta_p^2 = .22$; see Figure 5C). Follow-up paired-samples t-tests marginalised across electrodes indicated that this interaction was driven by a significantly larger LPP amplitudes for monetary than instructive feedback in the ICS subgroup ($t(8) = -8.25$, $p = .00002$), but not for the CB subgroup ($t(9) = -1.39$, $p = .20$). This indicates that LPP amplitudes differed between feed-

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

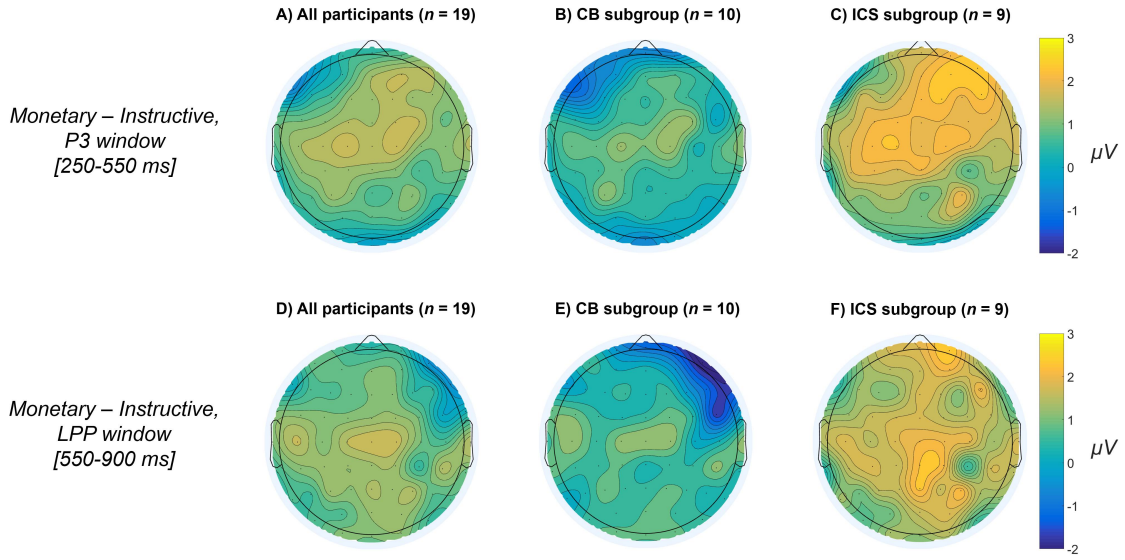

Figure 4. Scalp map of mean voltage difference between monetary and instructive feedback (in μV) for different model subgroups and analysis windows. Group-level voltage differences were largely driven by differences among participants in the Incentive-Compatible Switching (ICS) model subgroup, whereas voltage differences were small among participants in the Consistent Bayesian (CB) subgroup. Top row (A-C): P3 analysis window (250-550ms post-feedback). Bottom row (D-F): LPP analysis window (550-900ms post-feedback). Left column (A and D): Mean voltage difference pooled across all participants. Centre column (B and E): Mean voltage difference pooled across participants in the CB model subgroup. Right column (C and F): Mean voltage difference across participants in the ICS subgroup. For all scalp maps, voltages at missing electrodes have been reconstructed using spline interpolation for display purposes only.

back conditions only among participants who switched between a Bayesian strategy in the monetary feedback condition and a heuristic strategy in the instructive feedback condition; by contrast, there was no difference in LPP amplitudes for participants who employed a Bayesian strategy in both feedback conditions. No other main effects or interactions were significant, all $p > .10$.

Discussion

The present study assessed the effect of performance incentives on use of Bayesian versus heuristic strategies in a perceptual learning task. We found that, at a group level, participants' choices were more accurate when feedback was delivered in the form of monetary reinforcement than when it was delivered as instructive directives. Similarly, group-level results suggested differences between monetary and instructive feedback conditions in neural encoding of feedback in three ERP components: the P3, FRN, and the LPP. Critically, however, subsequent analyses informed by computational model comparison revealed that group-level behavioural and neural differences were, in fact, driven almost entirely by participants who switched between a Bayesian and a heuristic strategy according to feedback

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

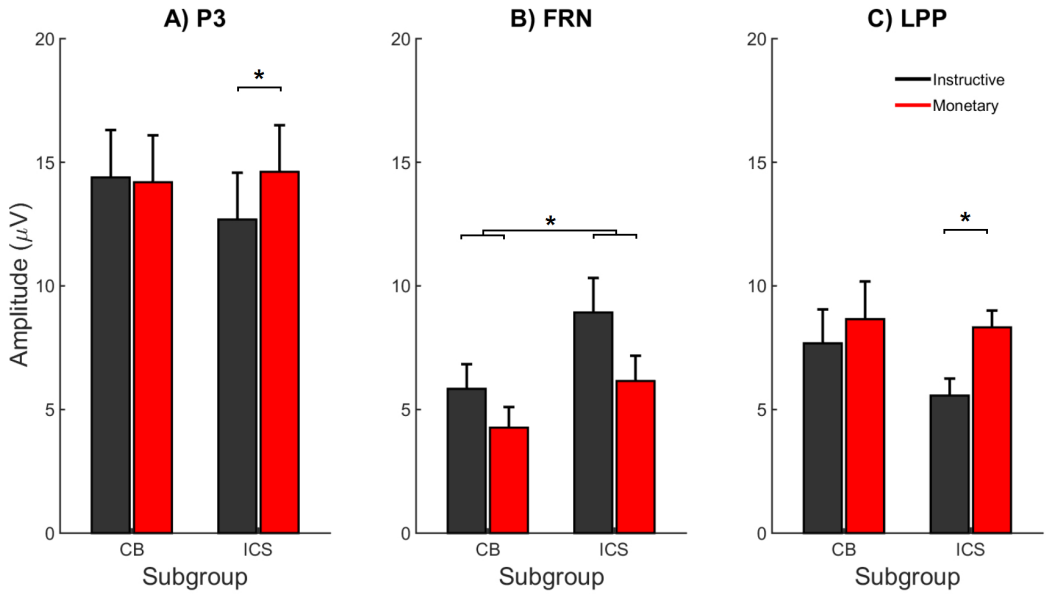

Figure 5. ERP component amplitudes as a function of feedback condition and behavioural subgroup (red: monetary feedback; black: instructive feedback). Error bars represent the standard error of the mean. * indicates p <.05. (A) P3 amplitude marginalised across electrodes AFz, Fz, FCz, Cz, and CPz. We observed a significant interaction between feedback condition and model subgroup, such that monetary feedback was associated with larger P3 components than instructive feedback for the ICS subgroup, but not for the CB subgroup. (B) FRN amplitude marginalised across electrodes AFz, Fz, FCz, Cz, and CPz. FRNs were significantly larger in the instructive feedback condition than the monetary feedback condition across both model subgroups. (C) LPP amplitude marginalised across electrodes Cz, CP, and Pz. We observed a significant interaction between feedback condition and model subgroup, such that monetary feedback was associated with larger LPP components than instructive feedback for the ICS subgroup, but not for the CB subgroup.

condition. In participants who consistently applied a Bayesian strategy in both feedback conditions, we observed no behavioural differences between monetary and instructive feedback.

Using a model-based cluster analysis based on formal comparison of computational cognitive models, we identified two distinct participant subgroups: a Consistent Bayesian (CB) subgroup, and an Incentive-Compatible Switching (ICS) subgroup. These two subgroups were associated with two models corresponding to qualitatively distinct behavioural strategies. The CB subgroup comprised participants best fit by a model assuming a Bayesian inference strategy in both the monetary and the instructive feedback condition. Conversely, participants in the ICS subgroup made choices more consistent with strategic switching between Bayesian inference in the monetary feedback condition and a WSLS heuristic strategy in the instructive feedback condition. This behaviour was consistent with incentive-compatible deployment of cognitive resources, at the cost of poorer performance in the instructive feedback condition. By contrast, participants in the CB subgroup used a Bayesian inference strategy in both feedback conditions, including when there was no monetary reward at

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

stake. Behavioural performance for the CB subgroup therefore did not differ between feedback conditions, and was consistently of a high standard overall. A control analysis revealed that model subgroup was unrelated to the order of feedback conditions, thereby ruling out a purely temporal switching effect.

ERP analyses revealed that, like behaviour, neural encoding of feedback also differed between CB and ICS subgroups. We assessed the effect of feedback condition and participant subgroup on the P3, FRN, and LPP: three ERP components associated with learning and processing of rewarding stimuli (Achtziger et al., 2015; Bennett et al., 2015; Frank et al., 2005; Yeung & Sanfey, 2004; Keil et al., 2002; Hajcak et al., 2009; Ito et al., 1998; Polich, 2007). This analysis showed an interaction of feedback condition and participant subgroup for the amplitudes of two components: the P3 and the LPP. This interaction was driven by differences between feedback conditions in the ICS subgroup only, such that monetary feedback elicited larger P3 and LPPs than instructive feedback. This was not the case for participants in the CB subgroup, who showed P3 and LPP components of similar amplitude in both feedback conditions, without any reduction for instructive feedback. Since only the ICS subgroup was associated with strategy-switching, this implicates the P3 and LPP as components which differentially encoded feedback depending on whether participants employed a Bayesian or a heuristic strategy. We also found a significant main effect of feedback condition for the FRN, indicating that across both participant subgroups, FRN amplitudes were larger for instructive than monetary feedback.

This differential neural encoding of feedback affords insight into the nature of feedback processing in Bayesian and heuristic strategies. In particular, the P3 has been linked in past research to the process of Bayesian belief updating (Bennett et al., 2015; Kolossa et al., 2015). It has been proposed that P3 amplitude indexes the magnitude of belief updates, possibly reflecting the deployment of working memory in the revision of prior beliefs (Kopp, 2008). Differences in feedback encoding between CB and ICS subgroups might therefore be interpreted as reflecting the differential engagement of a belief updating mechanism, since only Bayesian inference involves updating a full belief distribution. This is also in line with the proposal by Kok (1997) that P3 amplitude may reflect general cognitive effort, since Bayesian belief updating requires a greater expenditure of cognitive resources than a simple win-stay lose-switch heuristic.

Differential encoding of feedback in the LPP, by contrast, may reflect sensitivity to the reward valence of feedback. In tasks assessing encoding of affective stimuli, LPP amplitude has been associated with the affective salience of stimuli, such that both positively and negatively valenced stimuli elicited larger LPPs than neutral stimuli (Keil et al., 2002; Schupp et al., 2000). As such, one possible interpretation of LPP encoding differences in the present study is that the strategy-switching ICS group, but not the CB subgroup, perceived a difference in the emotional valence of monetary and instructive feedback. This may reflect a greater degree of reward sensitivity in the ICS subgroup than the CB subgroup, since it has previously been shown that reward processing may recruit different neural regions according to participants' reward sensitivity (Fröber & Dreisbach, 2014; Jimura et al., 2010).

Finally, we observed an overall effect of feedback condition on FRN amplitude, with a larger FRN for instructive compared to monetary feedback, but found that this effect did not interact with participant subgroup. This finding is in line with the hypothesis that FRN amplitude reflects a relatively automatic binary evaluation of stimulus valence (Yeung

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

& Sanfey, 2004), and may provide an electrophysiological index of affective components of feedback processing (Wiswede, Münte, Goschke, & Rüsseler, 2009). Our findings suggest that feedback value evaluation, as indexed by the FRN, was likely to have been equivalent in extent across all participants, independent of differences between participants in Bayesian versus heuristic task strategies. The smaller FRN elicited by monetary feedback in the present study may therefore reflect the greater overall hedonic value of monetary feedback relative to instructive feedback.

More broadly, it is important to note that the method of Bayesian model selection employed by the present study identifies which of a given set of computational models provides the most parsimonious account of behavioural data. Notably, therefore, this method does not provide information as to whether the best-fitting model within this set is also the best out of any possible model that might have been considered. For a task such as the perceptual learning task used in the present study, the space of possible models that might have been fit to the data is extremely large, and it was beyond the scope of the present study to exhaustively compare the fit of all possible learning models to participants' behaviour. Rather, the goal of model comparison in the present study was to assess the relative performance of two particular task strategies–Bayesian inference and a WSLS heuristic–that have been found to provide a good account of behaviour in our perceptual learning task by previous research (Bennett et al., 2015). The relative performance of different models was then used as a tool to make inferences regarding the effect of the feedback incentive manipulation on both behaviour and neural encoding of feedback.

More broadly, our results have bearing on the hypothesis that Bayesian inference represents a unifying principle of neural computation (the 'Bayesian brain' hypothesis; Knill & Pouget, 2004). This hypothesis has been applied successfully to domains including sensory coding and motor planning (Kording & Wolpert, 2004; Yuille & Kersten, 2006). However, one issue with applying Bayesian inference to higher-level judgement and decision making is that Bayesian inference is resource-intensive, and therefore computationally intractable for most real-world tasks(see e.g., Payzan-LeNestour & Bossaerts, 2011). Indeed, a wealth of evidence demonstrates that in many decision settings, humans fail to employ Bayesian strategies (e.g., Cassey, Hawkins, Donkin, & Brown, 2016; Gigerenzer & Goldstein, 1996). Moreover, even in cases where Bayesian inference is tenable, many individuals instead appear to rely on heuristic strategies (Bennett et al., 2015; Steyvers et al., 2009). Such evidence appears to challenge the suitability of simplistic Bayesian models to judgement and decision making. However, the results of the present study show that this impasse might be resolved by considering Bayesian models within a resource-rational framework that also takes cognitive resource limitations into account. Such an approach has been termed procedural rationality (Simon, 1976), or Type II rationality (Good, 1983). Our findings suggest that participants may select among Bayesian and heuristic behavioural strategies according to both associated outcomes and each strategy's processing costs (cf. Ortega & Braun, 2013). In computational terms these processing costs comprise both the computational expense of computing action policies, and the difficulty of learning (computational complexity versus sample complexity). This situates Bayesian models of cognition within an ecologically valid framework in which inference is constrained by the cognitive resource limitations of the human brain. From this perspective, we might conclude that ICS participants found the marginal value of employing a Bayesian over a heuristic strategy to be outweighed by the

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

cognitive costs of Bayesian inference in the instructive feedback condition.

In summary, using a model-based clustering analysis, we identified distinct subgroups of participants who appeared to use different combinations of Bayesian and heuristic strategies in a perceptual learning task. Incentive-compatible switching between Bayesian and heuristic strategies was associated with differences in performance between feedback conditions, as well as pronounced amplitude differences in ERP components linked to belief updating and affective salience processing. Overall, results suggest that motivational state may critically affect the use of Bayesian versus heuristic task strategies. This demonstrates the importance of embedding Bayesian models of cognition within a framework constrained by the cognitive resource limitations of biological agents. In addition, results suggest that individual differences in motivational state and reward sensitivity mediate the effect of incentives on task performance; as such, a one-size-fits-all approach to performance incentivisation in educational psychology or personnel economics is likely to be an oversimplification.

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

References

Achtziger, A., Alós-Ferrer, C., Hügelschäfer, S., & Steinhauser, M. (2015). Higher incentives can impair performance: neural evidence on reinforcement and rationality. *Social Cognitive and Affective Neuroscience*, *10*(11), 1477–1483.

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*(3), 610.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214.

Bennett, D., Murawski, C., & Bode, S. (2015). Single-trial event-related potential correlates of belief updating. *eNeuro*, *2*(5), ENEURO–0076.

Bode, S., Bennett, D., Stahl, J., & Murawski, C. (2014). Distributed patterns of event-related potentials predict subsequent ratings of abstract stimulus attributes. *PloS One*, *9*(10), e109070.

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*(3), 389.

Brydevall, M., Bennett, D., Murawski, C., & Bode, S. (2017). The neural encoding of information prediction errors during non-instrumental information seeking. *bioRxiv*. Retrieved from `http://www.biorxiv.org/content/early/2017/08/22/179507` doi: 10.1101/179507

Cassey, P., Hawkins, G. E., Donkin, C., & Brown, S. D. (2016). Using alien coins to test whether simple inference is bayesian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(3), 497.

Charness, G., & Levin, D. (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *The American Economic Review*, *95*(4), 1300–1309.

Chater, N., & Oaksford, M. (2008). *The probabilistic mind: Prospects for bayesian cognitive science*. OUP Oxford.

Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature*, *34*(2), 669–700.

Delorme, A., & Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Fraley, C., & Raftery, A. E. (1998). How many clusters? which clustering method? answers via model-based cluster analysis. *The Computer Journal*, *41*(8), 578–588.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501.

Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214.

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 530–547.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, *103*(4), 650.

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological Review*, *109*(1), 75.

Good, I. J. (1983). *Good thinking: The foundations of probability and its applications*. U of Minnesota Press.

Hajcak, G., Dunning, J. P., & Foti, D. (2009). Motivated and controlled attention to emotion: time-course of the late positive potential. *Clinical Neurophysiology*, *120*(3), 505–510.

Hidi, S. (2016). Revisiting the role of rewards in motivation and learning: implications of neuroscientific research. *Educational Psychology Review*, *28*(1), 61–93.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887.

Jimura, K., Locke, H. S., & Braver, T. S. (2010). Prefrontal cortex mediation of cognitive enhancement in rewarding motivational contexts. *Proceedings of the National Academy of Sciences*, *107*(19), 8871–8876.

Keil, A., Bradley, M. M., Hauk, O., Rockstroh, B., Elbert, T., & Lang, P. J. (2002). Large-scale neural correlates of affective picture processing. *Psychophysiology*, *39*(5), 641–649.

Knill, D. C., & Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719.

Kok, A. (1997). Event-related-potential (erp) reflections of mental resources: a review and synthesis. *Biological Psychology*, *45*(1), 19–56.

Kolossa, A., Kopp, B., & Fingscheidt, T. (2015). A computational analysis of the neural bases of bayesian inference. *Neuroimage*, *106*, 222–237.

Kopp, B. (2008). The p300 component of the event-related brain potential and bayes' theorem. In *Cognitive sciences at the leading edge* (pp. 87–96). Nova Science Publishers New York, NY.

Kording, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, *427*(6971), 244.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, *90*(5), 1346–1361.

Lopez-Calderon, J., & Luck, S. J. (2014). Erplab: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*.

Marković, D., Gläscher, J., Bossaerts, P., O'Doherty, J., & Kiebel, S. J. (2015). Modeling the evolution of beliefs using an attentional focus mechanism. *PLoS computational biology*, *11*(10), e1004558.

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*.

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality*, *60*(1), 117–174.

Moravec, H. P. (1988). Sensor fusion in certainty grids for mobile robots. *AI Magazine*, *9*(2), 61.

bioRxiv preprint doi: https://doi.org/10.1101/183665; this version posted September 1, 2017. 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-NC-ND 4.0 International license.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, *30*(37), 12366–12378.

Ortega, P. A., & Braun, D. A. (2013). Thermodynamics as a theory of decision-making with information-processing costs. In *Proceedings of the royal society a* (Vol. 469, p. 20120683).

Payzan-LeNestour, E., & Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology*, *7*(1), e1001048.

Pitz, G. F., & Sachs, N. J. (1984). Judgment and decision: Theory and application. *Annual Review of Psychology*, *35*(1), 139–164.

Polich, J. (2007). Updating p300: an integrative theory of p3a and p3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, *58*, 527–535.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2), 257–261.

Simon, H. A. (1976). From substantive to procedural rationality. In *25 years of economic theory* (pp. 65–86). Springer.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in Cognitive Science*, *7*(2), 351–367.

Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., & Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage*, *46*(4), 1004–1017.

Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, *53*(3), 168–179.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Cognitive Psychology*, *185*(4157), 1124–1131.

Wiswede, D., Münte, T. F., Goschke, T., & Rüsseler, J. (2009). Modulation of the error-related negativity by induction of short-term negative affect. *Neuropsychologia*, *47*(1), 83–90.

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258–6264.

Yuille, A., & Kersten, D. (2006). Vision as bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308.