Pre-registration Plan: Neural mechanisms in information sampling and belief updating with the Beads task

Christina Dimitriadou & Nicholas Furl

Study Information

1.1 Description

The present study investigates the mechanisms and neural underpinnings of optimal stopping problems. Optimal stopping problems are a subfield of probability theory; their basic framework is that the decision-maker or the observer, is involved in a process that evolves over time and includes some randomness. The goal of the observer is to sample as much evidence as is needed to minimise loss and maximise reward. When sampling evidence from the outside world, information is not abundant. In most cases information is unreliable or very noisy. Thus, in optimal stopping problems observers must commit to a decision and update their beliefs under high levels of uncertainty. The Beads task exemplifies how observers generate and update beliefs when uncertainty is high (Furl & Averbeck, 2011; Averbeck, 2015).

The Beads task is an information sampling task, in which individuals are presented with sequences of beads drawn from one of two possible urns (Fig 1A). Most common variants of the Beads task include monetary rewards when the individual correctly identifies the majority bead colour of the urn, and loses for incorrect responses. On each new draw in a sequence, individuals may choose to sample information (draw choice) up to *n* number of times before choosing to name the urn the beads are drawn from (urn choice; Fig 1B). The measure of interest is the number of draws before choosing an urn (draws-to-decision).

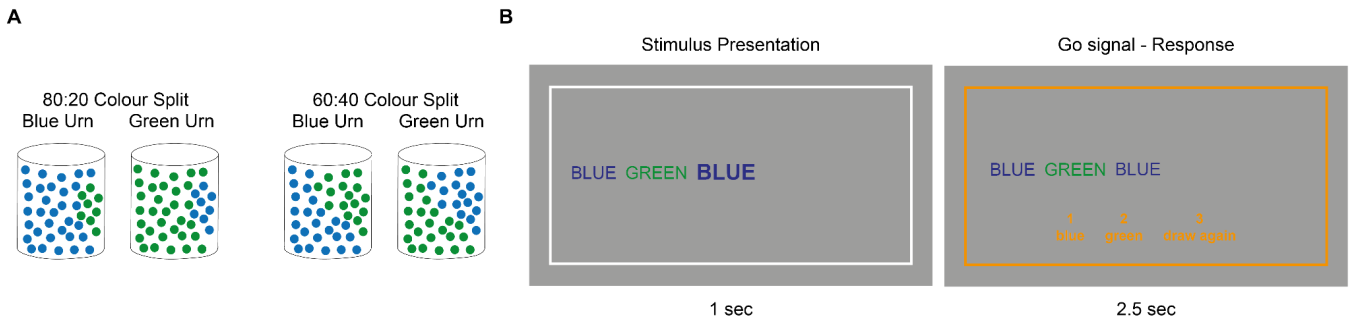

Figure 1. Description of the Beads task. **A)** Colour-split proportions of beads in the urns and experimental conditions. In the 80:20 colour-split condition, the blue urn contains 80% blue beads and 20% green beads, whilst the green urn contains 80% green beads and 20% blue beads. The same rule applies to the 60:40 condition. When a trial starts, participants know the colour-split and they are asked to infer the urn (blue or green) by drawing beads, one-by-one. **B)** Trial example. On every trial, participants are presented with a bead at the centre of the screen drawn from a hidden urn, whilst the inner rectange is white (for 1 sec). When the rectangle turns orange they can respond. They can either attempt to choose the majority colour of the hidden urn, or they can choose to draw again (sample more information). If they choose to draw again, the rectange turns white again, the previous bead is moved to the left side of the screen and the new bead appears at the centre of the screen. Participants are allowed to draw again up to nine times (to view all ten beads of that sequence) before choosing the urn that the balls are drawn from (blue or green). If they choose the correct urn, they win 10 credits, if they choose the incorrect urn they lose 10 credits, and they have a penalty of 0.25 credits for every time they choose to draw again. In this trial example, the hypothetical participant chose to draw again two times, thus viewed three beads drawn from the hidden urn.

At the neural level, the P300 component (Wang et al., 2015; Twomey et al., 2015) and beta oscillations (Fisher et al., 2018; Yaple et al., 2018) have been linked to decision-making and reward processing. Beta oscillations have also been linked to movement preparation (i.e., in decision-making tasks, beta power reflects motor responses; Kaiser et al., 2007). More recent studies suggest a more dynamic, gradual involvement of beta power as evidence is accumulated and the motor plan is continuously updated until a decision is made (Spitzer & Haegens, 2017). Even though most findings come from perceptual decision-making studies (Nunez et al., 2017) and are often modelled in the drift-diffusion framework (Bitzer et al., 2014), another line of work using the Beads task assessed relationships between probabilistic contingencies and the amplitude of frontal and parietal P300 amplitude (Kopp et al., 2016). The authors found that prior probabilities and likelihoods were associated with distinct neural responses, confirming that Bayesian inference is coded in the human brain.

Furthermore, Furl & Averbeck (2011) used a similar variation of the Beads task as shown in Figure 1, with fMRI. The authors reported larger neural responses when participants stopped the sequence and chose one of the two urns (urn choices), versus when they chose to draw again (draw choices) in a network including the anterior cingulate, striatum, insula and parietal cortex. Responses in the parietal cortex were stronger for subjects with increased draws-to-decision behaviour. These findings demonstrate that optimal stopping tasks can provide insight into the neural underpinnings of complex decision-making.

Optimal stopping problems and the Beads task can be modelled using a Bayesian ideal observer. Ideal observers perform the task in an optimal way (in the Bayesian sense), thus, allowing us to fit the model to human behaviour to quantify when is the optimal time to stop sampling information and commit to a decision. On a draw-by-draw basis, the model evaluates the expected values of drawing more beads versus choosing an urn (action values) until the value for choosing one of the two urns becomes smaller than the value for drawing again. Previous research on optimal stopping problems has used ideal observers in partially observable Markov decision processes (POMDP; Van de Wouw et al., 2021; Costa & Averbeck, 2015; Furl & Averbeck, 2011), and reported an undersampling bias for human participants compared to the ideal observer model.

In the present study, using the Beads task we will investigate this undersampling bias by comparing the amount of information sampled by human participants against the optimal choices made by a Bayesian ideal observer model in a partially observable Markov decision process (see Design and Analysis plan). Participants were performing the task while we recorded their EEG signals, thus, we will further investigate relationships between computations obtained by the model with underlying neural activity.

1.2 Hypotheses

At the behavioural level, based on previous findings (Furl & Averbeck, 2011) we hypothesise an effect of colour probability condition (colour probability 0.8 vs 0.6) on the number of draws and accuracy. In comparing agent types (human vs Bayesian ideal observer), we expect a main effect of agent type and colour probability and an interaction between agent type and colour probability.

At the neural level we expect differences in amplitude of the frontal and parietal P300 evoked component between colour probability conditions (probability 0.8 vs probability 0.6). We expect larger frontal and parietal P300 for trials at the 0.8 probability condition than for trials at the 0.6 probability condition. We further expect larger P300 amplitude over parietal versus frontal sites for urn choices than for draw choices. Regarding beta oscillations, we expect a relationship between beta power, probability conditions, draw choices and urn choices. We hypothesise slower beta oscillations (<20Hz) for draw choices and faster beta oscillations (>20Hz) for urn choices. We also expect larger amplitude in beta power response at the 0.8 probability condition.

Furl & Averbeck (2011) reported individual differences in parietal responses and the number of draws leading to choosing one of the urns. Based on this finding we hypothesise that the amplitude of the parietal P300 component and beta frequency will be predictive of participants' number of draws and of model action values.

Design Plan

2.1 Stimuli

The stimuli used in the experiment consisted of the words “blue” displayed in blue colour (RGB values: 30 70 13) and “green”, displayed in green colour (RGB values: 0 140 54).

2.2 Study design

2.2.1 Bayesian model

The Bayesian model will first be used as an ideal observer to evaluate human performance. This version of the model is parameter-free, meaning that parameters, such as cost to sample or cost for being wrong will be fixed. The ideal observer model balances the expected value for sampling more information (drawing more beads) against the expected reward of choosing an urn. In essence, for each draw of a given sequence, the model computes action values for the two urns and for drawing-again using the cost for choosing the incorrect urn and the cost for sampling more beads together with the probability that is tied to each of the urns. Once the action value for choosing one of the two urns exceeds the action value for drawing again, choosing an urn becomes the optimal choice and the evidence sampling stops. To run the ideal observer model we will use the same sequences as the participants used, therefore, performance of the model will be estimated for the sequences of every participant separately.

A second, parameterised model will be used that will be fit to each participant’s behavioural choices (responses to each draw on a given sequence) and sequences. In this model, we will let a parameter vary (i.e., cost to sample). Furl & Averbeck (2011), in two separate model fitting analyses, allowed either cost for being wrong or cost to sample to float. They reported that changing either parameter changed the model's draws-to-decision behaviour. However, in their formal analysis the authors decided to allow only cost for being wrong to vary as this parameter also varied in the experiment (the cost was either $0 loss or $10 loss), whilst cost to sample was kept constant at $0.25. In our experiment both cost for being wrong and cost to sample were kept constant throughout the experiment, therefore, allowing either

parameter to change, should bring the model's behaviour closer to the behaviour of the participants. As opposed to the ideal observer, this version of the model is regarded as a computational theory of optimal performance and will be used to look at ERP responses (P300) and oscillations (beta power) that relate to model parameters (action values).

2.2.2 Behavioural data

2x2 repeated-measure ANOVAs will be conducted with agent type as the first factor (human vs ideal observer) and colour probability condition (0.8 vs 0.6) as the second factor to compare performances in the number of draws (draws-to-decision) and accuracy. Note that in this set of repeated measures designs, human participants are paired with their corresponding instance of the ideal observer model, which means that in each pair, human and model were presented with the same sequences. Draws-to-decision will be calculated by simply averaging the number of draws for the two agent types and for probability conditions, and accuracy will be calculated by taking the proportion of correct responses (i.e., urns with the majority colour in the sequence).

2.2.3 EEG data

At the neural level, using SPM12 we will segment the continuous raw signal into epochs per colour probability condition (0.8 vs 0.6) and choices (draw choice vs urn choice) to obtain Event Related Potential (ERP) averages. This will result in four averaged ERPs: 0.8-draw choice, 0.8-urn choice, 0.6-draw choice and 0.6-urn choice. Time-Frequency (TF) analysis will also be conducted with a focus on peri-stimulus time courses (epochs) and beta power (13-30Hz). Time-Frequency Representations (TFRs) in the beta band will be averaged in the same way as the ERPs. Both for ERPs and TFRs, differences in probability conditions and choices will be contrasted and converted into 3D images.

A mass univariate analysis in interpolated sensor space will be conducted at the second (group) level to compute statistically significant mean differences between colour probability conditions and choice types contrasts, on the amplitude of P300 ERP and beta power.

To determine relationships of behavioural measures and neural activity for urn choices we will run one-sample t-tests using participants' averaged ERPs for urn choice and their average number of draws as a covariate. This analysis also will be performed in interpolated sensor space using the mass univariate approach.

Lastly, we will investigate linear relationships between action values obtained by the parameterised model and the ERPs and TFRs for urn choices and draw choices (for the two probability conditions) on an epoch-by-epoch basis using linear regression at the individual (first) level. Each participant's beta weights obtained at the first level analysis will then be used at the group (second) level in a one-sample t-test, to determine statistically significant differences from zero. Note that the Analysis Plan section provides a more detailed description of all statistical and computational models in the study.

2.2.4 Experimental design

On every trial/sequence, participants could draw up to nine times before choosing the majority urn the beads were drawn from. Each stimulus/bead was presented for 1 second but stayed on screen during the response time as well, for up to 2.5 seconds or until a

response was given (see Figure 1). Draw choices led to the presentation of a new bead and the cost of 0.25 credits for sampling again, while all the previously presented beads were presented on the left of the screen. Urn choices led to the presentation of a rating screen, where participants were asked to rate how confident they are about the choice they made on a scale from 0 to 100. After confidence-rating they were presented with a feedback screen, with either a win of 10 credits if the urn choice was correct or loss of 10 credits if the urn choice was incorrect. The confidence screen was self-paced, whilst the feedback screen stayed on for 1 second. Trial/sequence duration was based on the number of times participants chose to draw, thus, it varied between ~5.5 seconds (if a participant chose an urn at the first draw) to ~38 seconds (if a participant drew nine times before choosing the urn). Every new trial/sequence started with a self-paced information screen, displaying the updated credit balance, and the probability condition of that trial (0.8 or 0.6).

Participants were presented with 52 trials/sequences in total, split into four blocks of 13 trials. Similarly to the trial duration, the total duration of the experiment was also based on the number of times participants chose to draw and varied between ~10 minutes (if a participant always chose an urn after the first draw on every trial) to ~35 minutes (if a participant drew nine times before choosing the urn on every trial). At the end of the experiment, credits were converted to GBP and participants obtained 3% of the total winnings (up to £5).

2.3 Randomization

Of the 52 total trials, half were of 0.8 probability condition and the other half of 0.6 condition. Within the conditions, half of the trials (13 trials) were majority blue urn trials and the other half were majority green urn trials. Probability conditions (and majority urns) were randomised across the four blocks.

2.4 Manipulation check

Given the complexity of the experimental design, before performing the actual experiment participants were given detailed instructions of the task, were asked to complete a quiz of nine task-related questions, and to complete four practice trials. By ensuring that participants understood the manipulations of the task we will be able to assess whether all individual participants show a P300 response and beta power averaged across conditions.

Sampling Plan

3.1 Existing data

3.1.1 Registration prior to analysing the data

Data collection for the Beads task started in October 2021 and was completed in the beginning of May 2022. The dataset of a pilot participant has been used as an example dataset to create preprocessing pipelines at both the behavioural and neural levels. This dataset will not be used in formal analyses.

3.2 Participant recruitment procedure

The experiment has been approved by the College Ethics Committee at Royal Holloway University of London. Eligibility requirements included no neurological or psychiatric disorder, normal or corrected vision, no history of epilepsy in the family, no brain injuries, no medical prescription. Participants were recruited through the SONA recruitment platform and flyer advertisements. Each participant was paid £12 for participation and received a reward of up to £5 based on task performance.

3.3 Sample size

We recruited 40 participants to participate in the Beads task. In cases of incomplete datasets participants will not be replaced, but will be just excluded if over 20% of trials are missing.

3.4 Sample size rationale

A previous neuroimaging work on a Beads task used a sample of N=18 (Furl & Averbeck, 2011). Our sample of 40 participants exceeds this number and will give 80% power to detect medium group differences (d=0.65) at $p<0.05$.

Variables

4.1 Manipulated variables

At the behavioural level, we manipulated the colour-split in the urns (colour probability 0.8 and probability 0.6). In the 0.8 probability condition, 80% of the beads in the urn are of one colour (e.g. blue) and 20% of the other colour (e.g. green) and vice versa. The same rule applies to the 0.6 probability condition. At the neural level, manipulated variables consist of colour probability and choice types (urn choice vs draw choice).

4.2 Measured variables

The variables that we will measure in the Beads task are the number of draws up until choosing an urn (draws-to-decision), accuracy, action values after model fitting, and magnitudes of ERPs and beta power.

Acquisition

5.1 Computer screen

Stimuli were presented on SONY CPD-E530, CRT 21” display type, 1800 x 1440 resolution and 78 Hz refresh rate. The distance between participant and screen was 75 cm with no constraints.

5.2 EEG hardware and acquisition settings

EEG activity was recorded using a BioSemi ActiveTwo system (BioSemi, Inc., Netherlands) with 64 pre-amplified active Ag-AgCl electrodes, two mastoid electrodes and four EOG electrodes (vertical and horizontal). The BioSemi Active-Two system has two more

electrodes, the common mode sense (CMS) active electrode and the drive right leg (DRL) passive electrode, which were used as online reference and ground electrodes, respectively. We used the standard 10-20 international electrode montage, and electrode impedance was kept below 5 kΩ at all times, for all participants. EEG activity was recorded using the ActiView acquisition system at 512 Hz sampling rate.

Pre-processing

6.1 General set-up

All EEG data preprocessing and analysis steps have been scripted in Matlab 2021a (MathWorks Inc., MA United States) and the SPM12 toolbox (Penny et al., 2011) using the pilot dataset mentioned above. The planned preprocessing steps are described below.

6.2 Data import

Raw data were saved and exported in .bdf format for analyses. For each participant, we recorded and saved four .bdf files, one for each experimental block. For memory efficiency, the files will be imported and preprocessed one by one until merging. All recorded channels (EEG, mastoids and EOG) will be imported.

6.3 Conversion

Raw .bdf files will be imported into SPM12, read using default settings and automatically converted into SPM MEEG objects.

6.4 Channel definition

Channel definition will involve selecting channels that will be used for further preprocessing. At this stage, mastoid channels will be excluded and only EEG and EOG channels will be selected for analysis. We will re-reference the EEG signal to the average of all channels, a method that requires excluding noisy channels from re-referencing. During data collection, per participant, a log with details on the quality of the EEG recordings was kept, which included information about technical issues and any channels that contained noise (such as noise caused by head swinging and swaying and was persistent throughout the raw recording). Re-referencing to the average of all channels will be performed using a custom script, thus using the data-quality log, if the dataset of a given participant has a noisy channel, it will be excluded from the re-referencing script.

6.5 Montage creation and re-referencing

Montage creation involves re-defining EOG channels and specifying a reference using the channel-definition process described above. Ocular activity was recorded using four EOG electrodes (two vertical and two horizontal electrodes). Montage creation will derive one vertical and one horizontal channel by subtracting pairs of channels that were located above and below the right eye (vertical) and lateral of each eye (horizontal), respectively. The montage will also subtract the average of all EEG channels from each channel for re-referencing. For this step, we will use a custom script created in the lab.

6.6 High-pass filtering

We will use a high pass filter to remove ultra-low frequencies close to DC. We will high pass filter the data before resampling, as high-amplitude baseline shifts in the signal will otherwise generate filtering artefacts at the edges of the recording. The high pass cutoff frequency will be set to 0.1 Hz.

6.7 Resampling

As the 512 Hz recorded sampling rate is unnecessarily high, we will downsample the data to 256 Hz, thereby reducing the file size by half.

6.8 Low-pass filtering

Low pass cutoff frequency for this step will be set to 30 Hz. This low-pass cutoff will be used in the ERP analyses. For TF analyses we will low pass filter the resampled MEEG object with a cutoff set to 110 Hz.

6.9 Epoching and baseline correction

The continuous EEG signal will be epoched based on probability conditions (0.8 and 0.6 probabilities) and on choice conditions (draw and urn choice), thus, there are four event-types of interest: 0.8-draw and 0.8-urn events, 0.6-draw and 0.6-urn events. To epoch the signal we will use a trial definition function programmed specifically for this experiment. The total number of epochs depends on the total number of draws per participant. Generally, we know that there will be 52 urn epochs (26 0.8-urn and 26 0.6-urn epochs) as there were 52 sequences, but the number of draw epochs will vary. Based on the behavioural aspect of previous literature (see Furl & Averbeck, 2011), we expect that there will be more 0.6-draw epochs than 0.8-draw epochs, as participants tend to draw more in the 0.6 probability condition. Peristimulus time (epoch duration) will start at 500 ms before stimulus onset and will end at 800 ms after stimulus onset (-500 ms to 800 ms; 1.3 second total epoch duration). Baseline correction will be performed automatically, using the default SPM12 setting.

Given that we will low pass filter two MEEG objects (one filtered at 30 Hz for ERP analyses and another one at 110 Hz for TF analyses), the epoching and baseline-correction described above is meant to be applied to the 30 Hz MEEG object. The 110 Hz MEEG object will be epoched in the same way, however, we will not apply baseline-correction, as this step will take place after time-frequency decomposition.

6.10 Merging

Up until this step, each of the four data files per participant will be preprocessed individually. For each participant, we will merge the four MEEG objects produced during the previous steps and all individual/block files will be deleted. MEEG objects for ERP analysis (low pass filtered at 30 Hz) and for TF analysis (low pass filtered at 110 HZ) will be merged.

6.11 Artefact rejection

To detect and exclude epochs that contain artefacts and/or noise we will use the SPM12 thresholding method to all EEG channels. The threshold cutoff will detect and reject epochs

in which the signal recorded at any of the channels exceeds 100 microvolts relative to pre-stimulus baseline. Only the merged MEEG object for ERP analyses will undergo artefact rejection.

6.12 Averaging

To produce ERPs for the four conditions, we will perform robust averaging of the merged-cleaned data. Epochs will be averaged using the robust-averaging algorithm, as it suppresses artefacts that were not detected during artefact rejection, without rejecting epochs or channels completely but just the contaminated parts.

Analysis Plan

7.1 Statistical models

7.1.1 Behavioural analysis

Behavioural data will be preprocessed in Matlab version 2021a with a custom script and saved in .csv and .mat files for further analyses. Behavioural data will be analysed in Matlab using the *anovan.m* function. We will run two separate sets of 2x2 repeated-measures models with agent (human vs ideal observer) and probability (0.8 vs 0.6) as factors, one set on draws-to-decision and the second set on accuracy as dependent variables. Note that for these ANOVAs, we will use the output from the ideal observer model and not from the (second) parameterised model.

7.1.2 Ideal observer model

We will use the Bayesian ideal observer to model participants' behaviour by inputting as stimuli to the model the same sequences shown to each participant. In detail, the task will be modelled as a Bayesian finite-horizon Markov Decision Process with partially observable states (POMDP), based on previous studies (Averbeck, 2015; Furl & Averbeck, 2011). The model is designed to compute utilities u (i.e., the expected future rewards), and action values for a set of actions a (i.e., blue urn, green urn, draw again), for a given state s, at time t

$$u_t(s_t) = max_{a \in A_{s_t}} \{Q(s_t, a)\},$$

where action value Q is:

$$Q(s_t, a) = r(s_t, a) + C(s_t, a) + \sum_{j \in S} p(j|s_t, a) u_{t+1}(j).$$

Utility u of the state s at time t denotes the value of the best action A which depends on: the value of reward $r(s_t, a)$ acquired if action a is chosen, the cost-to-sample parameter $C(s_t, a)$, and the summation of probability outcomes j taken over subsequent states s at time $t + 1$. The term inside the brackets is used for the computation of action values for each possible action. To compute utilities, the model uses the max operator as it assumes a reward maximising agent, and so it chooses the action with the largest value.

For each draw, we will use the total number of beads that the participant has seen (n_d), the number of blue beads (n_b) and the probability for drawing again from the majority colour urn ($q = 0.8 \; or \; q = 0.6$) to compute the probabilities of each urn. In this example, the probability that we are drawing from the blue urn p_b is

$$p_b = \left[1 + \left(\frac{q}{1-q}\right)^{(n_d - 2n_b)}\right]^{-1}.$$

Similarly, $1 - p_b$ is the probability of the green urn. To estimate the values of choosing either the blue or the green urn, the cost for being correct and the cost for being incorrect will be used. In the beads task, the reward for being correct was set to 10 credits, the cost for being incorrect was -10 credits and the cost-to-sample for each new draw was set to -0.25 credits. A given state s consists of the number of draws n_d and the number of the majority colour beads (for this example, the blue urn n_b is used), and is written as $s_t = \{n_d, n_b\}$. The reward value for choosing action $a = blue$ is

$$r_b(s_t, a = blue) = C_e p_g + C_c p_b,$$

where C_e is the cost for being incorrect, C_c is the cost for being correct and p_b is the probability of the blue urn and p_g of the green urn. Similarly the value for the green urn, is computed as $r_g(s_t, a = green) = C_e p_b + C_c p_g$. The values for choosing either the blue or the green urn will also be regarded as $Q_b(s_t, a = blue)$ and $Q_g(s_t, a = green)$ when referencing the action values of the two urns. For drawing again, the action value is based on the cost-to-sample parameter and expected value of future states and is given by

$$Q_d(s_t, a = draw) = C_s + \sum_{j \in S} p(j|s_t, a) u_{t+1}(j),$$

where C_s is the cost for drawing again (cost-to-sample). If a participant chooses to draw again, the draw is either a blue or a green bead, so the subsequent state is $s_{t+1} = n_d + 1, \; n_b + 1$ if they draw a blue bead, or it is $s_{t+1} = n_d + 1, \; n_b$ otherwise.

7.1.3 Model fitting

The parameterised model will be fit to each participant's choices. We will include a free cost parameter (i.e., cost-to-sample) and by optimising this parameter the model will be able to predict individual participants' draw choices for each sequence that they were presented with. Thus, the aim here is to use the model to compute action values similar to each participant's values for every draw in each sequence and probability condition.

Other cost parameters of the model such as, cost for being incorrect and cost for being correct will be included as fixed parameters, therefore, in the fitting procedure, action values

for choosing one of the two urns will be estimated in the same way as in the ideal observer version (see previous section).

As the parameterised model fits a given participant's responses after every new draw to the model's action values, these values will be given by the Softmax function. For example, for blue urn choices, computing action probabilities from action values, will be given by

$$m_b = \left[\frac{e^{Q_b(s_t, a=blue)}}{\Sigma_{i \epsilon b,g,d} e^{Q_i(s_t, a=blue)}} \right].$$

To optimise the cost parameter we will use the simplex search algorithm in MATLAB (fminsearch). Using this method, we will first search for the cost-to-sample value that minimises the negative log-likelihood ($-\ ll$). Because $-\ ll$ is a loss function, the lower the value, the better the model fits participants' choice data. This will be given by

$$ll = -\Sigma_T \Sigma_{n_c} log\left(\Sigma_{i \epsilon b,g,d} D_i m_i\right).$$

The ll value will computed as a sum, taken over all T sequences and choices in each n_c sequence. D_i is set to 1 when the participant chooses action i, or it is set to 0 otherwise. Action values Q will be obtained after fitting the model to each participant's choices.

7.1.4 EEG analysis

We will first analyse the data in the time domain (ERPs) to examine evoked effects with a focus on the P300 component for all channels, in interpolated sensor space using Random Field Theory to correct for multiple statistical comparisons. To examine induced effects and oscillatory power we will further conduct Time-Frequency (TF) Analysis with a focus on beta power. Similarly to the evoked analysis it will be conducted for all channels, in interpolated sensor space using Random Field Theory. In addition, we will run two more analysis sets. The first set will investigate individual differences in relationships between participants' urn choice ERPs, beta power, and behavioural number of draws. The second analysis set will investigate draw-by-draw relationships between ERPs, beta power and action values computed by the parameterized model. Detailed steps of all analyses are listed below.

7.1.4.1 Analysis of evoked responses

Contrast calculation

After the last step of preprocessing (robust condition averaging) we will calculate difference contrasts for probability (0.6 versus 0.8) and choice (urn choice versus draw choice) conditions and their interaction (probability x choice) contrast.

Conversion to images

Analysis in interpolated sensor space using the mass univariate approach requires creating 3D scalp x time images by projecting sensor locations into a plane, and performing linear interpolation between them into a 32 x 32 pixel grid, tiled over all timepoints. Thus, before running scalp x time statistics, we will convert the contrasts computed at the previous step

into three-dimensional images. Note that for each of our EEG related analysis conducted in SPM12, conversion to images is always the last step of individual (or first) level analysis. All statistics will be calculated at the group (or second) level across all participants.

Mass univariate analysis

We will perform one-sample t-tests at the group-level in sensor space to examine whether difference contrasts and interaction images of averaged ERPs created at the individual level, differ from zero for all participants. This analysis will be conducted in three steps: model specification, where we will specify a design matrix for the one-sample t-test; model estimation, which will estimate the model parameters using the Restricted Maximum Likelihood (ReML) algorithm; and contrast specification which will produce scalp x time tables of statistics and statistical parametric maps showing significant clusters where conditions reliably differ.

7.1.4.2 Time-Frequency analysis

Wavelet estimation

For wavelet estimation we will use each participant's MEEG SPM object that will be low-pass filtered at 110 Hz, epoched without baseline-correction and merged. In detail, we will use Morlet wavelets to decompose each trial into power and phase across all channels, peristimulus time and frequencies. We will estimate frequencies from 1Hz to 55 Hz with seven wavelet cycles. This procedure will produce two files, one for power and one for phase; however, only the power file will be of interest in our analysis. All subsequent steps refer to analysis using only the power file.

Baseline rescaling

We will scale the power MEEG object produced in the previous step with a log transform as high frequency changes tend to be smaller than low frequency changes. We will use the log ratio option (LogR) and set the baseline time-window as -500 to -50.

Averaging power over frequency

For each participant, we will average the rescaled MEEG dataset over beta frequency range only (13 Hz to 30 Hz). After this step, we will perform robust averaging of all epochs to suppress potential artefacts. Note that this step is also required before contrast calculation.

Contrast calculation

Contrast calculation of the averaged TF epochs will be performed in the same way as for the averaged ERPs. We will calculate contrasts of the difference in power for probability (0.6 versus 0.8) and choice (urn choice versus draw choice) conditions and a contrast of their interaction (probability x choice).

Conversion to images

We will convert the TF contrast MEEG objects into three-dimensional images in scalp x time mode in the same way as with the ERP contrasts.

Mass univariate analysis

As with the ERP contrast images, we will perform one-sample t-tests at the group-level in sensor space to examine whether difference contrasts and interaction images of averaged TFRs created at the individual level differ from zero. We will perform the exact same three steps for model estimation to produce tables of statistics and parametric maps showing significant clusters where conditions in beta power reliably differ.

7.1.4.3 Individual differences in behaviour and evoked responses

The individual differences analysis that we will perform is motivated by a similar analysis in Furl & Averbeck (2011) on fMRI data. The purpose of this analysis is to investigate significant relationships between averaged ERP responses to urn choices and individual differences in participants' average number of draws. This will be performed in sensor space for all the EEG channels using the Mass Univariate approach.

Contrast calculation and conversion to images

To compute urn choice contrasts for each participant, we will use the averaged ERPs MEEG object produced in preprocessing - averaging step. Here we will set to 1 the two conditions that relate to the urn choices (0.8-urn choice and 0.6-urn choice) and to 0 the other two conditions (0.8-draw choice and 0.6-draw choice). This will result in a contrast file of only responses to urn choices. We will then convert the contrast file into a three-dimensional scalp x time image over all EEG channels.

Mass univariate analysis

Using the mass univariate approach, we will run a second-level (group level) model where we will input the urn choice (converted to images) contrasts of all participants in a one-sample t-test using participants' average number of draws as a covariate. Therefore, at the contrast specification step, we will apply to the design matrix a contrast that estimates positive differences from zero for urn choice evoked responses and contrasts that estimate positive and negative linear relations of urn choice evoked responses with the number of draws (the covariate).

7.1.4.4 Individual differences in behaviour and beta power

This analysis will investigate significant relationships between averaged beta power for urn choices and individual differences in participants' average number of draws. It will be performed in sensor space for all the EEG channels using the Mass Univariate approach.

Averaging power over frequency

For each participant we will use the rescaled power MEEG object produced at the step of baseline rescaling. We will average the data over beta frequency range (13 Hz to 30 Hz) only, as the focus of this analysis is the relationship of beta power and behaviour. After averaging TF power over beta frequency, we will perform robust averaging to suppress any potential artefacts and to average over epochs.

Contrast calculation and conversion to images

To compute urn choice contrasts for each participant, the step is identical to contrast calculation of averaged ERPs. The resulting contrast file will then be converted into a three-dimensional scalp x time image over all EEG channels.

Mass univariate analysis

Mass Univariate analysis will be identical to the analysis of averaged ERPs (see Individual differences in behaviour and evoked responses sub-section) using the 3D contrast images produced at the previous step.

7.1.4.5 Association between evoked responses and action values

The last two analyses will investigate relationships between evoked responses, beta power and action values of the parameterised Bayesian model. This set of analyses is linked to the last hypothesis.

Computation of grand average

We will first compute a grand average across all participants using each participant's averaged MEEG object produced at the last step of pre-processing. Using this grand averaged file, we will select a specific range of time-points (i.e., around 300 ms) and channels that show large responses to draw and urn choice conditions.

Cropping individual epochs and averaging channels and samples dimensions

We will then crop each participant's epoched MEEG object (produced at the step of merging) around the time and channel range specified at the previous step. In essence, this step will require cropping all epochs for a specific number of channels and samples (time-points) selected using the grand average, extracting the data from the cropped MEEG object and saving it as a new dataset (.mat file) for analysis outside SPM12. The resulting cropped datasets will be three-dimensional (channels x samples/trial x trials). Next, we will average channels and samples per trial, so that we get one data-point for each trial to be associated with the model's action values at a draw-by-draw basis.

Computation of the difference in action values between choice options for colour probabilities

Before looking at linear associations between evoked responses and model action values, we will compute the difference between draw-by-draw action values for choosing one of the two urns and choosing to draw again. Note, that these values will come from the parameterised model, where we will use each participant's sequences and responses for each draw.

Linear regression (within participants)

To examine whether model action values can predict the P300 response, for each participant we will run linear regression analysis using the cropped evoked responses as the dependent variable and difference in action values as the regressor.

Second-level analysis (across participants)

As a last step, we will use the beta weights obtained at the linear regression analysis to test whether the linear relationship between evoked responses and model action values are significantly different from zero, in a one-sample t-test across participants.

7.1.4.6 Association between beta power and action values

Computation of grand average

We will compute a grand average across all participants using each participant's MEEG object produced in the time-frequency analysis (averaging of conditions subsection). Using this grand averaged file, we will select a specific range of time-points (i.e., around 300 ms) and channels (e.g., frontal and parietal electrodes) that show larger response to draw and urn choice conditions.

Averaging individual epochs over frequency and time

We will then average each participant's rescaled TF MEEG object over frequency with a focus on beta power range, and over time using the time range selected during the previous step. Lastly, we will extract the data from the averaged TF MEEG object and save it as a new dataset (in .mat files) for analysis outside SPM12. The channel and sample dimensions of cropped beta power datasets will be averaged for each trial similarly to ERP cropped datasets.

Linear regression (within participants)

To examine whether model action values can predict beta power for drawing again and for choosing an urn, for each participant, we will run linear regression analysis using the averaged individual TF epochs as the dependent variable and difference in action values as the regressor.

Second-level analysis (across participants)

Lastly, we will use the beta weights obtained at the linear regression analysis to test whether the linear relationship between frontal and parietal beta power and model action values significantly differ from zero, in a one-sample t-test across participants.

7.2 Inference criteria

The conventional $p < 0.05$ criteria will be used to test statistical significance in the ANOVA models and regressions. To adjust for multiple comparisons of the neural data, the mass univariate approach uses Random Field Theory (RFT) correction, a parametric approach to controlling for type I errors. A $p < 0.005$ criteria will be used as a cluster-defining threshold and for all significant clusters (at that threshold) we will use p values < 0.05.

7.3 Data exclusion

Data from participants that failed to understand or follow the instructions of the study will be excluded.This will be accomplished by looking at the proportion of correct responses. If a given participant has proportion-correct less than 55% (close to chance), then the data will be excluded from further analysis. Additionally, we will exclude datasets with >20% missing trials due to technical issues and/or EEG equipment malfunction.

7.4 Missing Data

Ideally, each participant shall have 2x26 trials for analyses. Datasets with up to 20% missing trials will be included and missing data will be ignored. However, datasets that exceed the 20% cutoff in missing trials, will be excluded from analyses.

Other

8.1 References

Averbeck, B. B. (2015). Theory of choice in bandit, information sampling and foraging tasks. *PLoS computational biology*, *11*(3), e1004164.

Bitzer, S., Park, H., Blankenburg, F., & Kiebel, S. J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a Bayesian model. *Frontiers in human neuroscience*, *8*, 102.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436.

Costa, V. D., & Averbeck, B. B. (2015). Frontal–parietal and limbic-striatal activity underlies information sampling in the best choice problem. *Cerebral cortex*, *25*(4), 972-982.

Fischer, A. G., Nigbur, R., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2018). Cortical beta power reflects decision dynamics and uncovers multiple facets of post-error adaptation. *Nature Communications*, *9*(1), 1-14.

Furl, N., & Averbeck, B. B. (2011). Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions. *Journal of Neuroscience*, *31*(48), 17572-17582.

Kaiser, J., Lennert, T., & Lutzenberger, W. (2007). Dynamics of oscillatory activity during auditory decision making. *Cerebral cortex*, *17*(10), 2258-2267.

Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2005). Applications of random field theory to electrophysiology. *Neuroscience letters*, *374*(3), 174-178.

Kopp, B., Seer, C., Lange, F., Kluytmans, A., Kolossa, A., Fingscheidt, T., & Hoijtink, H. (2016). P300 amplitude variations, prior probabilities, and likelihoods: A Bayesian ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(5), 911-928.

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of mathematical psychology*, *76*, 117-130.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (Eds.). (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Spitzer, B., & Haegens, S. (2017). Beyond the status quo: a role for beta oscillations in endogenous content (re) activation. *eneuro*, *4*(4).

Twomey, D. M., Murphy, P. R., Kelly, S. P., & O'Connell, R. G. (2015). The classic P300 encodes a build-to-threshold decision variable. *European journal of neuroscience*, *42*(1), 1636-1643.

van de Wouw, D. S., McKay, R., & Furl, N. (2021). Methodological Remarks Regarding Optimal Stopping Tasks and the Implications for Sampling Biases.

Wang, L., Zheng, J., Huang, S., & Sun, H. (2015). P300 and decision making under risk and ambiguity. *Computational intelligence and neuroscience*, *2015*.

Yaple, Z., Martinez-Saito, M., Novikov, N., Altukhov, D., Shestakova, A., & Klucharev, V. (2018). Power of feedback-induced beta oscillations reflect omission of rewards: evidence from an EEG gambling study. *Frontiers in neuroscience*, 776.