



# Breaking Deadlocks: Reward Probability and Spontaneous Preference Shape Voluntary Decisions and Electrophysiological Signals in Humans

Wojciech Zajkowski<sup>1</sup> · Dominik Krzemiński<sup>1</sup> · Jacopo Barone<sup>1</sup> · Lisa H. Evans<sup>1</sup> · Jiaxiang Zhang<sup>1</sup>

Accepted: 3 November 2020 / Published online: 30 November 2020  
© The Author(s) 2020

## Abstract

Choosing between equally valued options is a common conundrum, for which classical decision theories predicted a prolonged response time (RT). This contrasts with the notion that an optimal decision maker in a stable environment should make fast and random choices, as the outcomes are indifferent. Here, we characterize the neurocognitive processes underlying such voluntary decisions by integrating cognitive modelling of behavioral responses and EEG recordings in a probabilistic reward task. Human participants performed binary choices between pairs of unambiguous cues associated with identical reward probabilities at different levels. Higher reward probability accelerated RT, and participants chose one cue faster and more frequent over the other at each probability level. The behavioral effects on RT persisted in simple reactions to single cues. By using hierarchical Bayesian parameter estimation for an accumulator model, we showed that the probability and preference effects were independently associated with changes in the speed of evidence accumulation, but not with visual encoding or motor execution latencies. Time-resolved MVPA of EEG-evoked responses identified significant representations of reward certainty and preference as early as 120 ms after stimulus onset, with spatial relevance patterns maximal in middle central and parietal electrodes. Furthermore, EEG-informed computational modelling showed that the rate of change between N100 and P300 event-related potentials modulated accumulation rates on a trial-by-trial basis. Our findings suggest that reward probability and spontaneous preference collectively shape voluntary decisions between equal options, providing a mechanism to prevent indecision or random behavior.

**Keywords** Decision making · Reward probability · Preference · EEG · Cognitive modelling

## Introduction

Cognitive flexibility enables decision strategies to adapt to environmental and motivational needs (Schiebener and Brand 2015). One characteristic of this ability is that harder decisions often take longer. Evidence from neurophysiology (Gold and Shadlen 2001), neuroimaging (Heekeren et al.

2008), and modelling (Ratcliff and Smith 2004) suggest an evidence accumulation process for decision-making: information is accumulated over time, and a decision is made when the accumulated evidence reached a threshold (Gold and Shadlen 2007). This process can accommodate paradigms consisting of noisy stimuli (perceptual choices), as well as a rich variety of tasks with unambiguous stimuli (value-based (Pisauro et al. 2017) or memory-based choices (Ratcliff 1978)). For perceptual choices, evidence is derived from the sensory properties of the stimuli; for value or preference-based choices, it originates from internal value evaluation and comparison (Krajbich et al. 2012); while for memory-dependent choices, from sampling memory traces (Ratcliff 1978; Shadlen and Shohamy 2016). According to this framework, decision difficulty, and in turn response time (RT), is proportional to the relative difference in the evidence supporting each option, consistent with results from perceptual (Ditterich et al. 2003), value-based (Polanía

---

Equal contributions: Wojciech Zajkowski, Dominik Krzemiński

✉ Wojciech Zajkowski  
zajkowskiw@cardiff.ac.uk

✉ Jiaxiang Zhang  
zhangj73@cardiff.ac.uk

<sup>1</sup> Cardiff University Brain Research Imaging Centre, School of Psychology, Cardiff University, Cardiff, UK

et al. 2014; Oud et al. 2016), and memory-based decisions (Ratcliff and McKoon 2008).

Making difficult choices requires more evidence; hence, longer deliberation can be an advantageous decision strategy. Yet, scaling deliberation with difficulty is beneficial only to a certain point. What happens if decision difficulty reaches a tipping point with values of options being indistinguishable? In the hypothetical paradox of *Buridan's ass* (van Inwagen 1989), a donkey which cannot choose between two identical haystacks would, as a result of its indecision, starve to death. This view is consistent with the classical drift-diffusion model (DDM, Ratcliff and McKoon 2008), which encodes the relative difference of evidence in favor of two options as a single accumulation process between two absorbing boundaries. Such a model would predict a deadlock or indecision between two equal alternatives because there is zero difference in the mean evidence supporting each choice (e.g., two identical haystacks), and the decision process is dominated by noise accumulated over time, resulting in prolonged RT (Teodorescu et al. 2016) (but see Ratcliff et al. (2018) for a recent model modification that addresses this theoretical limitation).

On the other hand, economic analysis suggests that choices between equal alternatives should be made as fast as possible. The benefit of “rushing to decisions” comes from being able to relocate our cognitive resources elsewhere (Rustichini 2009). If evidence cannot bring us closer to a better choice, deliberative thinking becomes an expensive and unnecessary luxury. This effect can be modelled using stochastic decision models with multiple accumulators, each encoding the accumulated evidence in favor of one choice, such as the Linear Ballistic Accumulator model (Brown and Heathcote 2008) and the Leaky Competing Accumulator model (Usher and McClelland 2001; Bogacz et al. 2007). For those models, multiple accumulators compete against each other on the basis of multiple sources of evidence inputs, which by default eliminates the scenario of indecision between equal alternatives.

In reality, individuals can make timely choices between equally valued options. For example, in preference-based decisions, it took under 2 s for one to choose between two snack food stimuli that had similar valuations (Voigt et al. 2019). In both humans and non-human primates, higher reward magnitude facilitates RT in perceptual and value-based decisions between equal choices (Pirrone et al. 2018). Intuitively, Buridan's donkey would be motivated to make faster decisions if the haystacks are fresh, compared to when they are stale. This magnitude effect is in line with ecological incentives: high rewards may imply a resource-rich environment, for which one needs to exploit as early as possible; low rewards may imply a resource-poor environment in which it is worth waiting for a better option (Pirrone et al. 2018). Furthermore, if choices are

based purely on expected rewards, one may choose any of the equal-valued options with the same frequency, leading to random behavior. Nevertheless, previous studies (Zhang and Rowe 2015; Phillips et al. 2018) showed that in a sequence of voluntary action decisions, humans deviated from a random pattern of choice and exhibited low choice entropy across trials. A similar conclusion has been reached in consumer decisions, where brand loyalties are driven by seemingly irrational preferences (Wheeler 1974). These findings suggest a possible preference bias between equal options, which renders some options more likely to be chosen than others.

We focus on three issues that have been unresolved in previous research on choices between equal alternatives. First, we aim to explore the effect of reward probability on RT. We expect that, similar to magnitude (Teodorescu et al. 2016; Pirrone et al. 2018), higher reward probability accelerates RTs. This prediction is not trivial, since probability and magnitude can have different effects on behavior. For example, Young et al. (2014) showed that magnitude discounting follows a power law, while probability is discounted hyperbolically. Unlike magnitude, probability has an upper bound at 100%, which acts in a qualitatively distinct way on behavior (Tversky and Kahneman 1989). We expect this increase in speed to be non-linear, with choices between two certain (100% probability) options being disproportionately faster compared to choices between two uncertain ones.

Second, in the evidence accumulation framework, both the rate of the accumulation and the non-decision time can influence a model's prediction of reaction time, the former encoding the strength of evidence and the latter reflecting the latencies of visual encoding and motor execution. During perceptual learning, the accumulation rate increases along with behavioral improvements (Jia et al. 2018), while the non-decision time remains unchanged in the late stage of training (Zhang and Rowe 2014). Furthermore, the accumulation rate is associated with the individual differences in working memory (Schmiedek et al. 2007) and attention (Nunez et al. 2017), while the non-decision time is faster in individuals with higher diffusion MRI-derived neurite density in the corticospinal tract, the primary motor output pathway (Karahan et al. 2019). Recent research showed that both parameters can be influenced by reward magnitude (Wagner et al. 2020), and the current study will examine further whether reward probability and preference influence the two model parameters.

Third, we aim to describe the macroscopic pattern of brain activities associated with differences in behavior: it is temporal evolution and relation to model-derived parameters. Functional imaging studies have localized the mesocorticolimbic dopaminergic network to be involved in both reward certainty and preference processing (Tobler

et al. 2007; Abler et al. 2009), but little is known about how these relate to global activations across the scalp. Pinpointing when EEG activity diverges between conditions and assessing whether these differences are transient or sustained can further inform our computational model, giving deeper insight into the cognitive underpinnings of the decision process.

Here, we address these questions by combining advanced computational modelling and EEG in a probabilistic reward task. Participants memorized six unambiguous cues associated with three levels of reward probability, a certain reward level (i.e., 100%) and two levels of uncertain reward probabilities (80% and 20%). Participants made two alternative forced choices between cues with equal reward probability (Fig. 1). The inclusion of the 100% reward probability condition allowed us to investigate whether cues with definitive rewards are processed in a different manner than the uncertain cues (Esber and Haselgrove 2011). Additional task conditions involved binary decisions between cues with different reward probability (unequal trials) and unitary responses to single cues (single-option trials). This design enabled us to focus on the neurocognitive processes underlying choices between equal options, while participants maintained a clear understanding of cue values for rational decisions between unequal options.

We first examine how reward probability influences behavior and whether a preference bias between equal options is present. We then fit an accumulator model of decision-making (Brown and Heathcote 2008) to the behavioral performance across reward probability levels. Posterior group parameters from hierarchical Bayesian

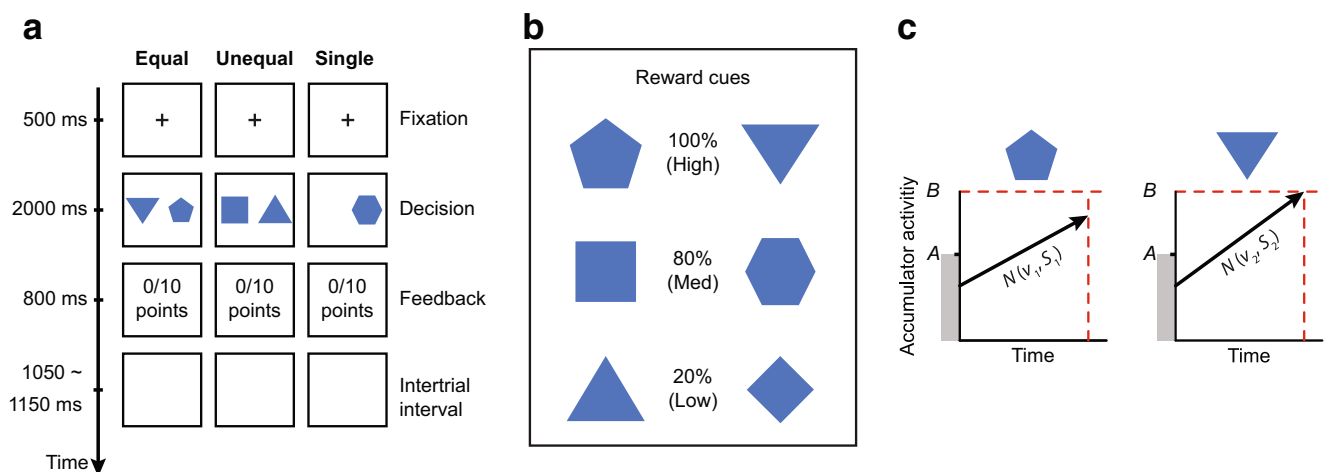
model fitting procedure were used to infer whether the behavioral effects were driven by evidence accumulation or non-decisional components of the process. EEG data were analyzed with time-resolved multivariate pattern classification for decoding spatiotemporal representations of reward probability and preference. To establish a link between the decision process and its EEG signatures, we integrated behavioral and EEG data into a joint hierarchical Bayesian model and tested the hypothesis that electrophysiological activity reflects trial-by-trial changes in the speed of evidence accumulation for decisions (Twomey et al. 2015).

We demonstrate that reward probability and spontaneous preference independently shape RTs and choices when deciding between equal alternatives. These behavioral effects affect the decision process and evoke a distinct electrophysiological pattern. Together, our findings contribute to the understanding of how decision deadlocks between two equally probable rewards can be overcome.

## Materials and Methods

### Participants

Twenty-three healthy participants were recruited from Cardiff University School of Psychology participant panel (20 females; age range 19–32, mean age 22.7 years; 22 right-handed). All participants had normal or corrected-to-normal vision, and none reported a history of neurological or psychiatric illness. Written consent was obtained



**Fig. 1** **a** Experimental paradigm of the probabilistic reward task. Participants were instructed to decide between two reward cues (equal and unequal trials) or respond to a single cue (single-option trials). **b** A total of six reward cues were randomly assigned to three levels of reward probability (100%, 80%, or 20%). **c** Exemplar time course of the linear ballistic accumulator (LBA) model for equal choices. On each trial, the LBA assumes that evidence for two options are

accumulated linearly and independently over time in two accumulators. The accumulation rate is sampled from a normal distribution with mean  $v$  and standard deviation  $S$ . The starting point of the accumulation process is sampled from a uniform distribution between 0 and  $A$ . The accumulation process terminates once the accumulated evidence first reaches a threshold  $B$ , and a corresponding decision is made by the winning accumulator

from all participants. The study was approved by the Cardiff University School of Psychology Research Ethics Committee.

## Apparatus

The experiment was conducted in a dedicated EEG testing room. A computer was used to control visual stimulus delivery and record behavioral responses. Visual stimuli were presented on a 24-inch LED monitor (ASUS VG248) with a resolution of  $1920 \times 1080$  pixels and a refresh rate of 60 Hz, located approximately 100 cm in front of participants. Participants' responses were collected from a response box (NATA technologies). The experiment was written in Matlab (Mathworks; RRID: SCR\_001622) and used the Psychophysics Toolbox Version 3 extensions (Kleiner et al. 2007).

## Experimental Design

All participants performed a decision-making task with probabilistic rewards during EEG recording (Fig. 1a). Before the task, participants memorized 6 unambiguous cues represented by different symbols and their associated probabilities of receiving a reward (Fig. 1b; see Procedure). All the cues had the same color (RGB = 246, 242, 92) on a black background (100% contrast). Each cue was mapped onto one of the three reward probability levels: *high* (a reward probability of 100%, i.e., always rewarded), *medium* (a reward probability of 80%), and *low* (a reward probability of 20%); and hence, there were two different cues associated with each reward probability.

Participants were instructed to maximize the total accumulated reward in the decision-making task. The task contained three types of trials: equal, unequal, and single option. On an equal trial, two different cues with the same reward probability appeared on the left and right sides of a central fixation point (e.g., 100% vs. 100%, 80% vs. 80%, or 20% vs. 20%). On an unequal trial, two cues with different reward probability levels appeared on both sides of the central fixation point (e.g., 100% vs. 20%, 100% vs. 80%, or 80% vs. 20%). On a single-option trial, one of the six cues appeared on either the left or right side of the fixation point. In equal and unequal trials, participants chose the left or right cue via button presses with the right-hand index and middle fingers. In single-option trials, participants responded to which side the single cue was presented (i.e., left or right). In all trials, the reward was operationalized as 10 virtual “game points” that did not have any tangible value. The probability of receiving the reward in a trial was either 100%, 80%, or 20%, which was determined by the chosen cue. It is worth noting that, in equal trials, participants' decisions did not actually affect the probability

of receiving the reward because both options had equal reward probability. In single-option trials, if participants chose the wrong side with no cue presented (0.1% across all single-option trials), no reward was given. Feedback for rewarded (a “10 points” text message on the screen) or not rewarded (blank screen) choices was given after each trial. The total game points awarded were presented at the bottom of the screen throughout the experiment.

## Procedure

Each experimental session comprised 640 trials, which were divided into 4 blocks of 160 trials. Participants took short breaks between blocks and after every 40 trials within a block. The mapping between the six reward cues and three levels of reward probability was randomized across participants. During breaks, the cues-reward mappings were explicitly presented on the screen (Fig. 1b), and the participants could take as much time as they needed to memorize them. After the first two blocks, all the cues were remapped to different reward probabilities. For example, for the pair of two cues that were associated with 100% reward probability in the first and second blocks, one of the two cues would be associated with 80% reward probability in the third and fourth blocks, and the other associated with 20% reward probability. Participants were encouraged to memorize the altered cue-probability associations prior to the third block. This remapping procedure reduced the potential bias associated with specific cues. No explicit memory tests were performed.

Each block contained 64 equal trials (32 for 100% vs. 100%, 16 for 80% vs. 80%, and 16 for 20% vs. 20%); 64 unequal trials (32 for 80% vs. 20%, 16 for 100% vs. 80%, and 16 for 100% vs. 20%) and 32 single-option trials (16 for 100%, 8 for 80%, and 8 for 20%) at a randomized order. This design ensured the same number of trials with and without cues with the highest reward probability (100%). Note however that individual cues did not differ much in terms of frequency of occurrence: each 100% cue appeared 56 times, compared to 48 for each non-certain cue. This makes it unlikely that observed differences can be explained by occurrence frequency alone. Because two cues were bound to every probability level, different cue positions and combinations can result in the same reward probability pair (e.g., there are 4 possible combinations for 80% vs. 20% unequal trials). These combinations were counterbalanced across trials.

Each trial began with the presentation of a fixation point at the center of the screen for 500 ms. After the fixation period, in the equal and unequal trials, two reward cues appeared on the left and right sides of the screen with a horizontal distance of  $4.34^\circ$  from the fixation point. Both cues were vertically centered. In single-option

trials, only one reward cue appeared on one side of the screen, and the side of cue appearance was randomized and counterbalanced across trials. Cues were presented for a maximum of 2000 ms, during which participants were instructed to make a left or right button press. The cues disappeared as soon as a response was made, or the maximum duration was reached. The reaction time (RT) on each trial was measured from the cue onset to button press. Reward feedback was given 200 ms after the reward cue offset and lasted 800 ms, followed by a random intertrial interval uniformly distributed between 1050 and 1150 ms. As in our previous study (Zhang and Rowe 2014), if the participant failed to respond within 2000 ms or responded within 100 ms, no reward was given and a warning message “Too slow” or “too fast” was presented for 1500 ms.

### Behavioral Analysis

We excluded trials with RT faster than 200 ms (fast guesses). For each participant, trials with RTs longer than 2.5 standard deviations from the mean RT were also excluded from subsequent analysis. The discarded trials accounted for 1.5% of all trials.

We first analyzed the proportion of choices in equal trials to establish the existence of a preference bias. In the equal condition, by definition, there was no “correct” or “incorrect” response, since the cues had the same reward probability. For each pair of cues with the same reward probability, we defined the preferred cue as the one chosen more frequently than the other (non-preferred) in equal trials. The categorization of preferred and non-preferred cues was estimated separately between the first two and the last two blocks because of the cue-probability remapping after the first two blocks. At each level of reward probability, a preference bias was then quantified as the proportion of trials where the preferred cue was chosen. The preference bias had a lower bound of 50%, at which both cues were chosen with equal frequency.

In the unequal condition, we defined decision accuracy as the proportion of choosing the cue with higher reward probability, separately for each combination of reward probabilities (100% vs. 80%, 100% vs. 20%, and 80% vs. 20%). Two-tailed one-sample *t* tests compared the decision accuracy in the unequal condition against a chance level of 50%, which would indicate irrational decisions (i.e., both high and low reward cues were chosen in 50% of trials).

To determine how reward probability, preferences, and other experimental factors influence RT, we analyzed single-trial RT data with linear mixed-effects models (LMMs) using the *lme4* package (Bates et al. 2015) in R (RRID: SCR\_001905). The LMM is a hierarchical regression method that distinguishes between fixed and random effects Gueorguieva and Krystal (2004). LMMs take into account all single-trial data without

averaging across trials and offer better control of type 1 and type 2 errors than ANOVA (Baayen et al. 2008). Therefore, statistical inferences from LMMs are robust to experimental designs with unbalanced trials across conditions (Bagiella et al. 2000), which is an important feature suitable for the current study.

We designed two LMMs with different dependent variables and factors (Table 1). Model 1 analyzed the RTs from equal and single-option trials, including choice type (equal or single-option), reward probability (high, medium or low), cue remapping (before and after), preference (whether the chosen cue was preferred), and right-side bias (whether the chosen cue was on the right side of the screen) as factors. Right-side bias was included to control for spatial bias relating to preference for stimuli presented on the right or left side of the screen. For the unequal condition, because each trial had two cues with different levels of reward probability that cannot be directly compared with equal or single-option trials, the RTs were analyzed separately in model 2. Here, we used similar predictors with exception of probability, which was captured by two additional factors: the sum and the absolute difference of the two reward probabilities, as they both have been shown to affect choice behavior (Thaler 1991; Ballard et al. 2017; Teodorescu et al. 2016).

In all the LMMs, fixed effects structures included hypothesis-driven, design-relevant factors and their interactions, and individual participants were included as the source of random variance (random effect). We used a standard data-driven approach to identify the random effects structure justified by the experimental design, which resulted in good generalization performance (Barr et al. 2013). This approach starts with the maximal random effects structure (i.e., including all random slopes, intercepts and interactions) and systematically simplified it until the LMM reaches convergence. Table 1 lists the simplified random effects structures. The correlation structures of each fitted LMM were assessed to avoid overfitting (Matuschek et al. 2017).

### A Cognitive Model of Voluntary Decision-Making

We further analyzed the behavioral data using the Linear Ballistic Accumulator (LBA) model (Brown and Heathcote 2008). LBA model is a simplified implementation of a large family of sequential sampling models of decision-making (Ratcliff and Smith 2004; Bogacz et al. 2006; Gold and Shadlen 2007; Zhang 2012) which assumes an independent accumulation process for each choice option. Our model-based analysis has three stages. First, we fit a family of LBA models with various model complexity to the behavioral data of individual participants in equal trials. By identifying the best-fitting model, we infer how reward probability

**Table 1** The linear mixed-effects models of RT

	Model 1	Model 2
Dependent variables	RT	RT
Main effects	Reward probability Preference Cue-remapping Choice (equal of single-option) Right-bias	Sum of reward probability Difference of reward probability Preference Cue-remapping Right-bias
Interaction terms	Probability * choice Probability * preference Probability * cue-remapping Choice * preference Choice * cue-remapping Preference * cue-remapping Probability * choice * preference Probability * choice * cue-remapping Probability * cue-remapping * preference	Sum of reward probability * preference Difference of reward probability preference Sum of reward probability * cue-remapping Difference of reward probability * cue-remapping Preference * cue-remapping
Random effects (correlated slopes) and intercepts)	Reward probability preference Cue-remapping Choice Right-bias	Sum of reward probability Difference of reward probability Preference Cue-remapping Right-bias

Model 1 analyzed single-trial RT in equal and single-option trials. Model 2 analyzed single-trial RT in unequal trials. In both models, preference was a predictor indicating whether the preferred cue was selected in each trial. Cue-remapping was a predictor indicating whether each trial was before or after cue-probability remapping in the second half of each session. Right-bias indicated whether the cue on the right size of the screen was chosen in each trial, modeling a possible response bias

and preference modulated subcomponents of the evidence accumulation process during decision-making. Next, we simulate the best-fitted LBA model and examine whether model simulations are consistent with the experimental data in single-option and unequal conditions. This is a stringent test of model generalizability because the experimental data in single-option and unequal trials are unseen by the model-fitting procedure. Finally, we link the cognitive processes identified by the LBA model to brain activities by incorporating a trial-by-trial measure of EEG activity regressors into the best-fitted model (Cavanagh et al. 2011; Nunez et al. 2017, 2019).

The LBA model assumes that the decision of when and which to choose is governed by a “horse race” competition between two accumulators  $i \in \{1, 2\}$  that accumulate evidence over time supporting the two choice options (Fig. 1c). One accumulator is in favor of the preferred cue and the other of the non-preferred cue. The activations of the accumulators represent the accumulated evidence. At the beginning of each trial, the initial activations of the two accumulators are independently drawn from a uniform distribution between 0 and  $A$ . The activation of each accumulator then increases linearly over time,

and the speed of accumulation (i.e., accumulation rate) varies as a Gaussian random variable with mean  $v_i$  and standard deviation  $S_i$  across trials. The accumulation process terminates when the activation of any accumulator reaches a response threshold  $B$  ( $B > A$ ) and the choice corresponding to the winning accumulator is selected. The model prediction of RT (measured in seconds) is the sum of the duration of the accumulation process and a constant non-decision time  $T_{er}$ , with the latter accounting for the latency associated with other processes including stimulus encoding and action execution (Brown and Heathcote 2008; Nunez et al. 2019; Karahan et al. 2019).

### Model Parameter Estimation and Model Selection

LBA model has five key parameters: mean  $v$  and standard deviation  $S$  of the accumulation rate across trials, decision threshold  $B$ , starting point variability  $A$ , and non-decision time  $T_{er}$ . To accommodate the empirical data, one or more model parameters need to vary between conditions. We evaluated a total of 21 variants of the LBA model with different parameter constraints (Fig. 3a). First, the accumulation process may differ between the preferred

and non-preferred options, leading  $v$  or  $S$  to vary between accumulators (preferred, non-preferred). Second, reward probability could modulate the accumulation process or visuomotor latencies unrelated to decisions, leading to  $v$ ,  $S$  or  $T_{er}$  to vary between three levels of reward probability. Third, the decision threshold  $B$  and starting point  $A$  were fixed between conditions because the trial order was randomized, and we do not expect the participants to systematically vary their decision threshold before knowing the cues to be presented (Ratcliff and Smith 2004). Fourth, decision threshold  $B$  and starting point  $A$  were fixed across preference levels, since participants could not predict which cue would appear on which side of the screen. During model-fitting, the decision threshold was fixed at 3 as the scaling parameter (Brown and Heathcote 2008), and all the other parameters allowed to vary between participants. Theoretically, the scaling parameter can be set to an arbitrary value, which does not influence the parameter inference, as long as the priors of other parameters remain realistic, but with some constraints parameter estimation is easier to converge. Finally, because the participants showed behavioral differences between reward probability levels and between preferred/non-preferred choices, we only estimated realistic models: those with at least one parameter varied between reward probability levels ( $v$ ,  $S$ , or  $T_{er}$ ) and at least one parameter varied between accumulators ( $v$  or  $S$ ).

We use a hierarchical Bayesian model estimation procedure to fit each LBA model variant to individual participant's choices (the proportion of preferred and non-preferred choices) and RT distributions in equal trials. The hierarchical model assumes that model parameters at the individual-participant level are random samples drawn from group-level parameter distributions. Given the observed data, Bayesian model estimation uses Markov chain Monte Carlo (MCMC) methods to simultaneously estimate posterior parameter distributions at both the group level and the individual-participant level. The hierarchical Bayesian approach has been shown to be more robust in recovering model parameters than conventional maximum likelihood estimation (Jahfari et al. 2013; Zhang et al. 2016).

For group-level parameters ( $v$ ,  $S$ ,  $A$ , and  $T_{er}$ ), similar to previous studies (Annis et al. 2017), we used weakly informed priors for their means  $E(\cdot)$  and standard deviations  $std(\cdot)$ :

$$\begin{aligned} E(v) &\sim N(2.5, 1), & std(v) &\sim \gamma(1, 1), \\ E(S) &\sim N(1, 0.75), & std(S) &\sim \gamma(1, 1), \\ E(A) &\sim N(2.5, 1), & std(A) &\sim \gamma(1, 1), \\ E(T_{er}) &\sim N(0.5, 0.2), & std(A) &\sim \gamma(1, 1), \end{aligned} \quad (1)$$

where  $N$  represents a positive normal distribution (truncated at 0) with parameterized mean and standard deviation, and

$\gamma$  represents a gamma distribution with parameterized mean and standard deviation.

We used the *hBayesDM* package (Ahn et al. 2017) in R for the hierarchical implementation of the LBA model. For each of the 21 model variants, we generated four independent chains of 7500 samples from the joint posterior distribution of the model parameters using Hamiltonian Monte Carlo (HMC) sampling in Stan (Carpenter et al. 2017). HMC is an efficient method suitable for exploring high-dimensional joint probability distributions (Betancourt 2017). The initial 2500 samples were discarded as burn-in. To assess the convergence of the Markov chains, we calculated Gelman-Rubin convergence diagnostic  $\hat{R}$  of each model (Gelman et al. 1992) and used  $\hat{R} < 1.1$  as a stringent criterion of convergence (Annis et al. 2017). We compared the fitted LBA model variants using Bayesian leave-one-out information criterion (LOOIC). LOOIC evaluates the model fit while considering model complexity, with lower values of LOOIC indicating better out-of-sample model prediction performance (Vehtari et al. 2017).

## EEG Data Acquisition and Processing

EEG data were collected using a 32-channel Biosemi ActiveTwo device (BioSemi, Amsterdam). Due to technical issues, EEG data collection was not successful in two participants; and therefore, all EEG data analyses were performed on the remaining 21 participants. EEG electrodes were positioned at standard scalp locations from the International 10-20 system. Vertical and horizontal eye movements were recorded using bipolar electrooculogram (EOG) electrodes above and below the left eye as well as from the outer canthi. Additional electrodes were placed on the mastoid processes. EEG recordings (range DC-419 Hz; sampling rate 2048 Hz) were referenced to linked electrodes located midway between POz and PO3/PO4 respectively and re-referenced off-line to linked mastoids. Additional electrodes were placed on the mastoid processes. EEG (range DC-419 Hz; sampling rate 2048 Hz) was collected with respect to an active electrode (CMS, common mode sense) and a passive electrode (DRL, driven right leg), which were located midway between POz and PO3/PO4 respectively, to form a ground-like feedback loop.

EEG data were pre-processed using EEGLab toolbox 13.4.4b (Delorme and Makeig, 2004; RRID: SCR\_007292) in Matlab. The raw EEG data were high-pass filtered at 0.1 Hz, low-pass filtered at 100 Hz using Butterworth filters and downsampled to 250 Hz. An additional 50-Hz notch filter was used to remove main interference. We applied independent component analysis (ICA) to decompose continuous EEG data into 32 spatial components, using *runica* function from the EEGLab toolbox. Independent components reflecting eye movement artifacts were identified by the

linear correlation coefficients between the time courses of independent components and vertical and horizontal EOG recordings. Additional noise components were identified by visual inspection of the components' activities and scalp topographies. Artifactual components were discarded, and the remaining components were projected back to the data space.

After artifact rejection using ICA, the EEG data were low-pass filtered at 40 Hz and epoched from  $-400$  to  $1000$  ms, time-locked to the onset of the stimulus (i.e., reward cues) in each trial. Every epoch was baseline corrected by subtracting the mean signal from  $-100$  to  $0$  ms relative to the onset of reward cues.

### Multivariate Pattern Analysis

We use time-resolved Multi-Voxel Pattern Analysis (MVPA) on pre-processed, stimulus-locked EEG data to assess reward-specific and preference-specific information throughout the time course of a trial. In contrast to univariate ERP analysis, MVPA combines information represented across multiple electrodes, which has been shown to be sensitive in decoding information representation from multi-channel human electrophysiological data (Cichy et al. 2014; Dima et al. 2018).

We conduct three MVPA analyses to identify the latency and spatial distribution of the EEG multivariate information. The first to decode reward probability levels in equal choices (e.g., equal trials with two 100% reward cues versus equal trials with two 80% cues). The second to decode preferred versus non-preferred choices in equal trials. The third to decode between equal and single-option choices with the same reward probability (e.g., equal trials with two 100% cues versus single-option trials with a 100% cue).

Each analysis is formed as one or multiple binary classification problems, and the data feature for classification included EEG recordings from all 32 electrodes. In each analysis, at each sampled time point ( $-400$  to  $1000$  ms) and for each participant, we train linear support vector machines (SVM) (Garrett et al. 2003) using the 32-channel EEG data and calculate the mean classification accuracy following a stratified tenfold cross-validation procedure. In all MVPA, we include the EEG data from 400 ms before cue onset as a sanity check because one would not expect significant classification before the onset of reward cues.

In each cross-validation, 90% of the data issued as a training set, and the remaining 10% as a test set. In some analysis (e.g., equal trials with 100% cues versus equal trials with 80% cues), the number of samples belonging to the two classes is unbalanced in the training set. We use a data-driven over-sampling approach to generate synthetic instances for the minor class until the two classes had balanced samples (Zhang and Wang 2011). The synthetic

instances are generated from Gaussian distributions with the same mean and variance as in the original minority class data. Training set data were standardized with  $z$ -score normalization to have a standard normal distribution for each feature. The normalization parameters estimated from the training set were then applied separately to the test set to avoid overfitting. To reduce data dimensionality, we perform principal component analysis to the training set data and selected the number of components that explained over 99% of the variance in the training set. The test set data are projected to the same space with reduced dimensions by applying the eigenvectors of the chosen principal components. We then train SVM to distinguish between the two classes (i.e., conditions) and evaluate the classification accuracy using the test set data. The procedure is repeated ten times with different training and test sets, and the classification accuracies are averaged from the tenfold cross-validation. We use the SVM implementation in MATLAB Machine Learning and Statistics Toolbox. The trade-off between errors of the SVM on training data and margin maximization is set to 1.

To estimate the significance of the classification performance, we use two-tailed one-sample  $t$  test to compare classification accuracies across participants against the 50% chance level. To account for the number of statistical tests at multiple time points, we use cluster-based permutation (Maris and Oostenveld 2007) to control the family-wise error rate at the cluster level from 2000 permutations.

### Estimation of Single-Trial ERP Components

We estimate two ERP components from single-trial EEG data in equal trials: N100 and P300, which are subsequently used to inform cognitive modelling. The visual N100 is related to visual processing (Mangun and Hillyard 1991) and the P300 is related to evidence accumulation during decision making (Kelly and O'Connell 2013; Twomey et al. 2015).

To improve the signal-to-noise ratio of single-trial ERP estimates, we use a procedure similar to previous studies (Kayser and Tenke 2003; Parra et al. 2005; Nunez et al. 2019). For each participant, we first performed singular value decomposition (SVD) to the grand averaged ERP data across all trials from the same experimental condition. SVD decomposes the trial-averaged ERP data  $A_{k \times p}$  (where  $k$  is a number of channels and  $p$  is a number of time points) into independent principal components. Each component consists of a time series of that component and a weighing function of all channels, defining the spatial distribution (or spatial filter) of that component. Because the ERP waveform is the most dominant feature of the trial-averaged ERP data, the time course of the first principal component (i.e., the one that explains the most variance) represents a



cleaned trial-average ERP waveform (Nunez et al. 2019), and its weight vector provides an optimal spatial filter to detect the ERP waveforms across EEG channels. We then applied the spatial filter from the first principal component as a channel weighting function to single-trial EEG data to improve the signal-to-noise ratio.

The single-trial EEG data filtered with the SVD-based weighting function is then used to identify the peak latency and peak amplitude of the N100 and P300 components. For N100, we search for the peak negative amplitude in a window centered at the group-level N100 latency (112 ms) and started at 60 ms. The lower bound of the search window was determined by the evidence that the visual-onset latency is 60 ms in V1 (Schmolsky et al. 1998). For P300, we search for a peak positive amplitude in a window centered at the group-level P300 latency (324 ms). For both N100 and P300, the search window has a length of 104 ms, similar to a previous study (Nunez et al. 2019).

### EEG-Informed Cognitive Modelling

Recent studies showed that the variability of the P300 component closely relates to the rate of evidence accumulation during decision making (Twomey et al. 2015). We therefore extend the best-fitting LBA model with EEG-informed, single-trial regressors, which estimates the effect of trial-by-trial variability in EEG activity on the mean accumulation rate (Hawkins et al. 2015; Nunez et al. 2017).

The main regressor of interest is the slope of change between the N100 and P300 components, which is defined as the ratio of the P300–N100 peak-amplitude difference and the P300–N100 peak-latency difference in each equal trial. We also test four additional regressors from individual ERP components: P300 amplitude, P300 latency, N100 amplitude, and N100 latency. All the EEG regressors are obtained from the estimations of single-trial ERP components in equal-choice trials. To obtain a meaningful intercept, the regressors are mean-centered and rescaled to have a unit standard deviation.

Each EEG regressor is tested in a linear regression model, using the same Bayesian hierarchical model estimation procedure as in the behavioral modelling analyses. For each regression model, we assume that the mean accumulation rates of both accumulators  $v_1(t)$  and  $v_2(t)$  (i.e., the one in favor of the preferred option and the other one in favor of the non-preferred option) are influenced by the EEG regressor of interest on a trial-by-trial basis:

$$v_1(t) = \tilde{v}_1 + \beta \times EEG(t), \quad v_2(t) = \tilde{v}_2 + \beta \times EEG(t), \quad (2)$$

where  $t = 1, 2, 3, \dots$  represents the equal-choice trials, and  $\tilde{v}_1$  and  $\tilde{v}_2$  are the intercepts. The regression coefficient  $\beta$  represents the effect of EEG regressor on the mean accumulation rates.

The rationale of estimating an EEG regressor to the mean drift rate is twofold. First, this approach allows quantifying the trial-by-trial change over the intercept (i.e., the mean drift rate), independent of its trial-by-trial variability (parameter  $S$ ). Second, one would not expect the sensor-level EEG signal has sufficient spatial resolution to distinguish between the two accumulators encoding two options. Therefore, we estimated a single EEG regressor across both accumulators.

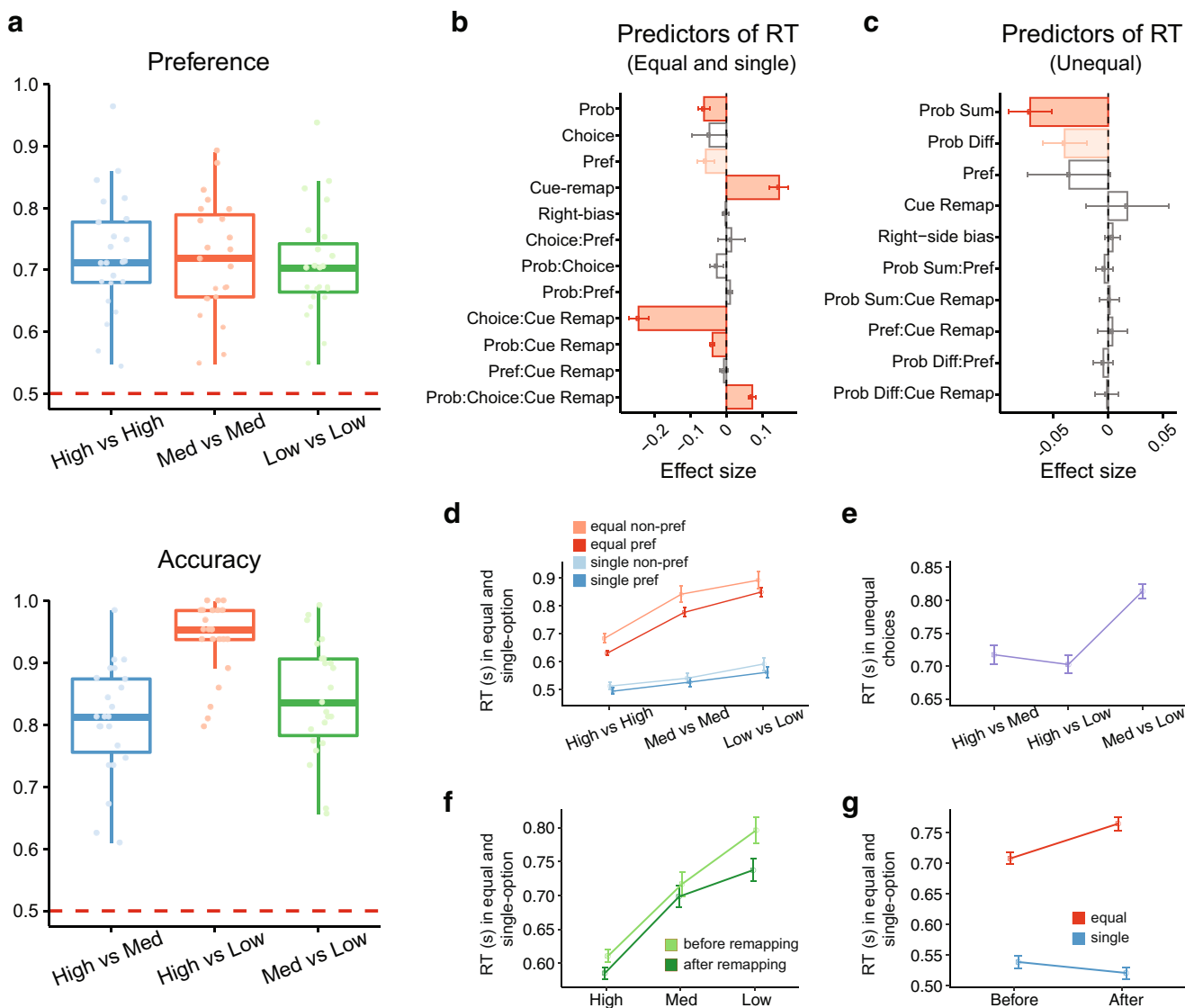
## Results

We examined the effects of reward probability and spontaneous preference on behavior and EEG activity during voluntary decisions. In a probabilistic reward task (Fig. 1b), participants chose between two options with the same reward probability (equal trials) at high (100%), medium (80%), or low (20%) levels. In two control conditions, participants made binary choices between options with different levels of reward probability (unequal trials) or responded to the location of a single reward cue (single-option trials). Below, we first report behavioral results. We then fit linear ballistic accumulator (LBA) models (Fig. 1c) to the choices and RT distributions of equal trials and infer about the underlying cognitive processes based on best-fitting model parameters. Next, we perform univariate and multivariate analyses of EEG data to identify spatiotemporal representations of reward probability and preference information as well as their time courses. We then extend the best-fitted LBA model with single-trial measures of EEG activity to test whether trial-to-trial variations in EEG data relates to the rate of evidence accumulation across trials.

### Behavioral Results

#### Choices

For each pair of cues with the same reward probability, we defined the preferred cue as the one chosen more frequently than the other (non-preferred) in the equal choice trials (see “Behavioral Analysis” in “Methods”). We found a strong preference bias (>50%) for choosing one reward cue over the other at each level of reward probability (Fig. 2a; high: 95%CI [0.682, 0.765]; medium: 95%CI [0.679, 0.759]; low: 95%CI [0.669, 0.745]). A repeated-measures ANOVA showed no significant difference in preference between reward probability levels ( $F(2, 44) = 0.2, p = 0.81$ ). Therefore, although the two options were associated with the same level of reward probability, participants did not make their choices randomly. We further used a linear mixed-effects model (LMM) to evaluate the preference



**Fig. 2** Behavioral results. **a** Preference bias across reward probability levels in equal trials (top) and decision accuracy across reward probability levels in unequal trials (bottom). **b** Linear mixed-effects model results for model 1 in Table 1. Dark red bars represent significant effects with  $p < 0.001$ . Light red bars represent significant effects with  $p < 0.05$ . Gray bars represent non-significant factors and interactions. Error bars represent standard errors across participants. **c** Linear mixed-effects model results for model 2 in Table 1. Significant effects

and interactions in RT from model 1 (Table 1) were presented separately for the following: reward probability and preference in equal and single-option trials (**d**), before and after cue-remapping at different reward probability levels (**f**), before and after cue-remapping in equal and single-option trials (**g**). Significant main effects in model 2 were presented in panel **e**. In panels **d–g**, error bars represent standard errors across participants

bias as a function of cue remapping (i.e., before vs. after the cue remapping halfway through each session) and trial order in each testing block. The preference bias was smaller after cue remapping (Supplementary Figure S1,  $\beta = -0.181$ , 95%CI  $[-0.01, -0.348]$ ,  $p < 0.03$ ), but was not influenced by trial order ( $\beta = 0.037$ , 95%CI  $[-0.170, 0.243]$ ,  $p = 0.73$ ). These results imply that, for a given set of cue-probability associations, the extent of preference bias did not significantly vary over time. Because the cue-probability mapping was randomized across participants and remapped within each session, the observed preference

bias is unlikely to be explained by a group-level preference toward any specific cue, but rather a spontaneous preference at the individual level. Additionally, to check if preference from first half of the experiment affected preference after remapping, we calculated the proportion of any cue being preferred in both sessions on a subject level. We found that preference was consistent only in 51.5% of cases, rendering no support for preference transfer after remapping (one-sided binomial-test  $p = 0.5$ , 95%CI =  $[0.361, 1]$ ).

In unequal choice trials, as expected, the cues with higher reward probability were chosen more often, as evidenced

by the above-chance decision accuracies in all conditions (Fig. 2a; high vs. medium:  $t(22) = 16.08$ , 95%CI [0.774, 1],  $p < 0.001$ ; high vs. low:  $t(22) = 23.31$ , 95%CI [0.862, 1],  $p < 0.001$ ; medium vs. low:  $t(22) = 20.97$ , 95%CI [0.834, 1],  $p < 0.001$ ; one-sample  $t$  test against the 0.5 chance level). A repeated-measures ANOVA showed significant differences in decision accuracy between reward probability levels ( $F(2, 44) = 28.17$ ,  $p < 0.001$ ). Post hoc pairwise comparison with Tukey's correction indicated that accuracy in the high vs. low probability condition (93.8%) was significantly higher than in the high vs. medium (84.3%) ( $t(44) = 5.267$ ,  $p < 0.001$ ) and the medium vs. low (80.7%) ( $t(44) = 7.265$ ,  $p < 0.001$ ). Similar to the analysis of preference, we used a LMM to evaluate decision accuracy in unequal trials as a function of cue remapping and trial order and found no significant associations (Supplementary Figure S1, cue remapping:  $\beta = 0.022$ , 95%CI [-0.235, 0.1901],  $p = 0.84$ ; trial order:  $\beta = 0.078$ , 95%CI [0.013, 0.169],  $p = 0.1$ ). These results suggested that participants memorized the cue-probability associations for rational choice behavior and maintained the decision accuracy throughout the experiment.

### Response Times

We used a LMM to quantify the influence of experimental factors on RTs in equal and single-option choices (Fig. 2b, Model 1 in Table 1). The fixed effects included reward probability, choice type (equal vs. single-option), preference (choosing the preferred vs. the non-preferred option), cue remapping, and their meaningful interactions (Fig. 2d–f). Participants were faster when choosing the preferred than the non-preferred option (Fig. 2d,  $\beta = -0.063$ , 95%CI [-0.027, -0.991],  $p < 0.05$ ) and RTs decreased as the reward probability increased ( $\beta = -0.101$ , 95%CI [-0.067, -0.135],  $p < 0.001$ ). The RT in equal choice trials was longer than that in single-option trials ( $\beta = -0.292$ , 95%CI [-0.201, -0.384],  $p < 0.001$ ). The effect of reward probability on RT was stronger in equal compared to single-option choices, supported by a significant interaction between the two main effects ( $\beta = 0.045$ , 95%CI [0.025, 0.066],  $p < 0.001$ ).

Participants had slower responses after memorizing a new set of cue-probability associations, indicated by a significant main effect in RT before and after cue remapping ( $\beta = 0.149$ , 95%CI [0.096, 0.201],  $p < 0.001$ ). The significant interaction between cue remapping and reward probability suggested that the increase in RT was more pronounced in trials with lower reward probability (Fig. 2f,  $\beta = -0.039$ , 95%CI [-0.051, -0.026],  $p < 0.001$ ). The interaction between cue remapping and choice type (Fig. 2g,  $\beta = -0.247$ , 95%CI [-0.192, -0.302],  $p < 0.001$ ) indicated that this pattern was mainly associated with equal

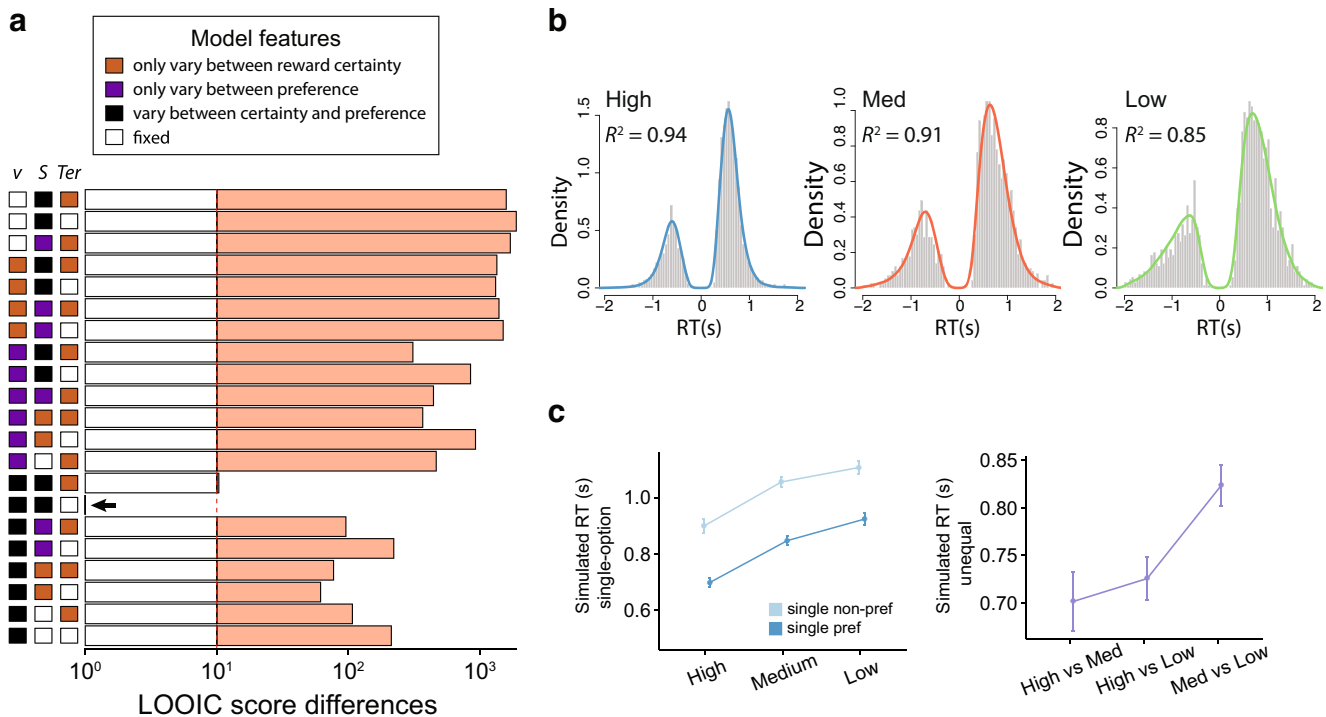
trials. Because evaluating reward probability of a cue was likely associated with additional cognitive load after cue remapping, the observed RT difference before and after cue remapping implies that participants evaluated both cues throughout the experimental session.

In a second LMM, we analyzed RTs in unequal trials (Model 2 in Table 1), including the sum and difference of the reward probability of two cues in each trial as fixed effects. The sum of two reward probabilities in unequal trials was negatively associated with RT (Fig. 2e,  $\beta = -0.071$ , 95%CI [-0.032, -0.110],  $p < 0.001$ ), consistent with previous studies that the total reward magnitude influences decision-making (Pirrone et al. 2018; Teodorescu et al. 2016). Additionally, the difference of two reward probabilities was also a significant predictor at a more lenient threshold ( $\beta = -0.028$ , 95%CI [-0.001, -0.055],  $p < 0.05$ ). No other effects or interactions reached significance, further solidifying that the cue-probability associations were well remembered in both halves of the experiment.

### Cognitive Modelling of Behavioral Data

To identify the cognitive processes that led to the observed behavioral differences, we compared 21 variants of the LBA model. The model variants differed systematically in their constraints on whether the rate of evidence accumulation and non-decision time could change between reward probability levels or preferred/non-preferred options. For each model variant, we used hierarchical Bayesian modelling with Markov chain Monte Carlo (MCMC) parameter estimation routine to estimate the posterior distributions of the model parameters, given the observed choice and RT distribution from individual participants (see “Model Parameter Estimation and Model Selection”). To identify the model with the best fit, we calculated the Bayesian LOOIC score for each model (Vehtari et al. 2017).

MCMC chains representing posterior parameter estimates in all the 21 model variants reached high levels of convergence (Gelman-Rubin convergence diagnostic  $\hat{R} \leq 1.02$  for all parameters in all models). The LOOIC scores suggested that the models with the mean accumulation rate varying between reward probability levels and between preference levels fitted the data better than others model variants. The best-fitting model (i.e., the one with the lowest LOOIC score, Fig. 3a) had fixed group-level non-decision time with the standard deviation of the accumulation rate varying between reward probability levels and preferred/non-preferred options. To evaluate the model fit to the empirical data in equal trials, we calculate the posterior prediction of the best-fitting model by averaging 100 iterations of model simulation using posterior parameter estimates. Averaging across multiple iterations reduces potential biases when sampling from posterior parameter



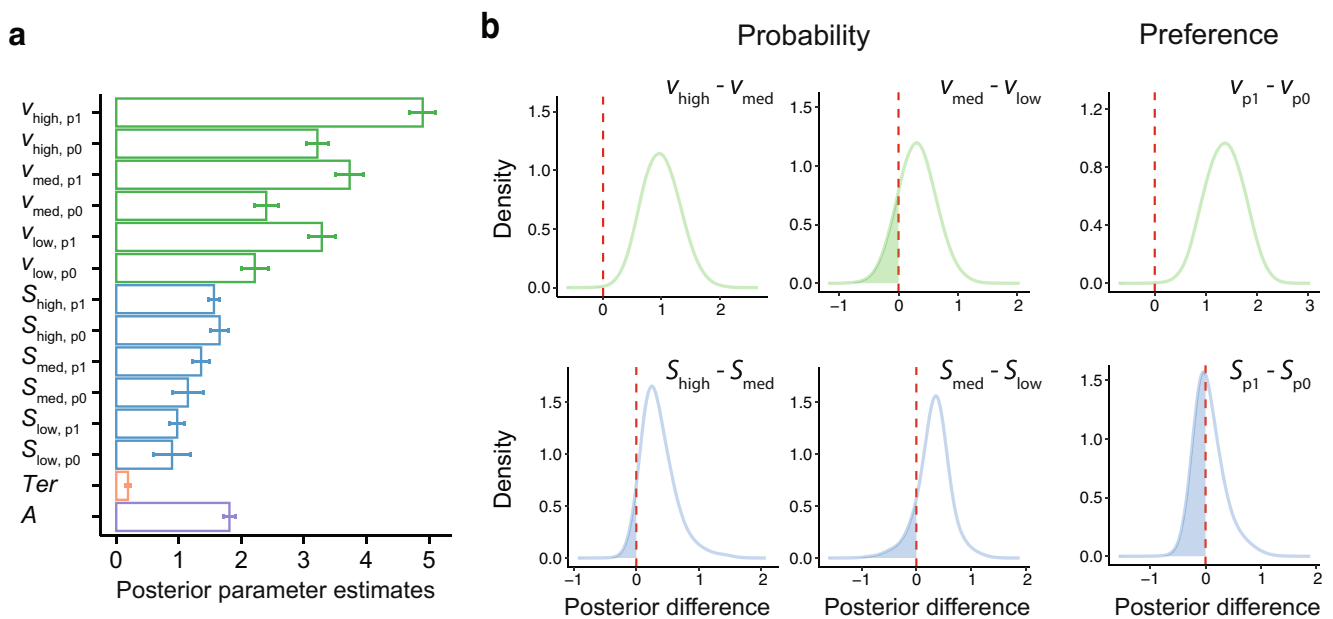
**Fig. 3** Model comparisons, model fits, and model simulations. **a** LOOIC scores of 21 LBA model variants. The LOOIC score differences between all models and the best model are plotted against corresponding model structures, which were illustrated on the left of the figure. The model structure specified how the mean accumulation rate  $v$ , the standard deviation  $S$  of the accumulation rate, and the non-decision time  $T_{er}$  could vary between conditions. A black-filled square indicated that the corresponding parameter could vary between reward probability levels and preferred/non-preferred options. An orange- or purple-filled square indicated that the corresponding parameter could only vary between reward probability levels or preferred/non-preferred options, respectively. Unfilled (white) squares indicated that

the parameter remained fixed between conditions. Bar color indicates whether the difference in LOOIC scores is considered substantial (over 10); white part of the bar corresponds to score up to 10, orange to the amount exceeding 10. The best model was shown with a LOOIC score difference of 0 (indicated by the red arrow). **b** Simulations of RTs in equal choices, generated from the posterior distribution of the best-fitted model for high (left), medium (middle), and low (right) reward probability levels. Histograms represent experimental data and density distributions represent model simulation from 100 iterations. Negative values represent RTs for non-preferred choices. **c** Simulation of RTs in single-option (left) and unequal (right) choices from 100 iterations. Error bars represent standard errors across participants

estimates. Each of the 100 iterations generates simulated behavioral responses (i.e., RTs and choices) of individual participants, with the same number of trials per condition as in the actual experiment. There was a good agreement between the observed data and the model simulations across reward probability levels and choice preferences (Fig. 3b).

We use Bayesian inference to analyze the posterior distributions of group-level model parameters (Bayarri and Berger 2004). To evaluate if a parameter varies substantially between any two conditions, we calculate the proportion of posterior samples in which the parameter value for one condition was greater than the other. To test if a parameter differs from a threshold value, we calculate the proportion of the posteriors greater or smaller than the threshold. To avoid confusion, we use  $p$  to refer to classical frequentist  $p$  values, and  $P_{p|D}$  to refer to Bayesian inference results based on the proportion of posteriors supporting the testing hypothesis, given the observed data.

For the best-fitting model (Fig. 4a), we compared the posterior estimates of the group-level parameters between conditions (Fig. 4b and c). We found strong evidence for choices with high reward probability to have higher mean ( $v$ ) and standard deviation ( $S$ ) of the accumulation rate than choices with medium ( $v_{high} > v_{medium} : P_{p|D} = 0.999$ ;  $S_{high} > S_{medium} : P_{p|D} = 0.954$ ) or low medium ( $v_{high} > v_{low} : P_{p|D} = 1$ ;  $S_{high} > S_{low} : P_{p|D} > 0.999$ ) reward probability. The mean and standard deviation of accumulation rates between choices with medium and low reward probabilities were inconclusive ( $v_{medium} > v_{low} : P_{p|D} = 0.839$ ;  $S_{medium} > S_{low} : P_{p|D} = 0.877$ ). Furthermore, there was also strong evidence for a higher mean accumulation rate for the preferred than the non-preferred options ( $P_{p|D} = 0.999$ ), and no evidence for a difference in the standard deviation of the accumulation rate ( $P_{p|D} = 0.532$ ). These results supported the claim that preferred and certain (100%) cues were recalled and



**Fig. 4** Posterior model parameters and inferences. **a** Group-level LBA model parameters of the best-fitting model: means of accumulation rates ( $v$ , green), standard deviations of accumulation rates ( $S$ , blue), non-decision time ( $T_{er}$ , orange) and starting point ( $A$ , purple). Error bars represent standard deviations of posterior distributions of parameter values. The means and standard deviations of accumulation rates were shown separately for each reward probability level

(high, medium, and low) and accumulator (p1, preferred option; p0, non-preferred option). **b** Differences of posterior parameter estimates across probability levels (left and middle columns) and preference levels (right column). The proportion of posterior difference distributions above 0 suggested higher parameter values for higher probability level or more preferred options

processed faster than non-preferred cues. Certain cues were also associated with more variable accumulation rate. Model comparisons further suggested that the latencies of early visual encoding and motor execution were not influenced by reward probability nor preference as the models with varying non-decision time parameter did not fit the data as well.

Next, we evaluated whether the best-fitting model could reproduce qualitative RT patterns in the single-option and unequal choices, which were unseen by the parameter estimation procedure. This allows us to evaluate whether the model that fits to the equal choice data can also characterize behavioral patterns in other conditions. For unequal choices, two accumulators representing two cues with different reward probability levels compete to reach the decision threshold, with their parameters set to the posterior estimates from the fitted LBA model. For single-option choices, a single accumulator is set to reach to the decision threshold. Similar to the simulation of equal choices, we average the predicted behavioral responses of unequal and single-option choices for each participant from 100 iterations of simulation. Each iteration contains the same number of trials as in the experiment.

For unequal trials, the simulated RT showed similar patterns to the observed data, in which choosing between medium and low probability cues led to the longest

RT (Fig. 3c). For single-option choices, similar to the observed data, higher reward probability and preferred cues were associated with faster RT in simulation. However, simulated RT in single-option choices was longer than the experimental data, suggesting that simple reactions to a single cue may engage distinct cognitive processes beyond the current model.

### EEG Results

We focused our EEG analysis on equal trials (with additional control analysis on EEG data from single-option trials) because both reward probability and preference bias played major roles in shaping the behavioral performance of that condition.

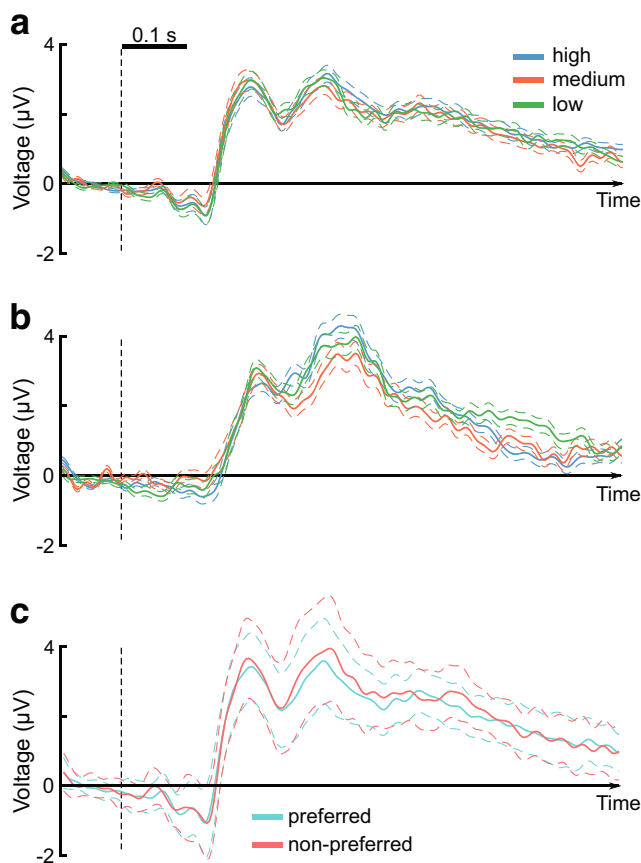
### Event-Related Potentials

We examine univariate differences in evoked responses between conditions in single EEG electrodes. For each participant, trial-averaged ERPs are calculated from epochs of equal or single-option choices, with epochs time-locked to reward cue onset. For both equal and single-option conditions, we test for differences in ERPs between three levels of reward probability using a one-way repeated-measures ANOVA. Furthermore, we test for differences in

ERPs between preferred and non-preferred choices in equal trials using a paired *t* test. We perform statistical tests on all electrodes and all time points. Cluster-based permutation tests (2000 iterations with maximum statistics) are used to correct for multiple comparisons across electrodes and time points (Maris and Oostenveld 2007).

Different reward probability levels produced similar grand-average ERP waveforms during equal (Fig. 5a) and single-option (Fig. 5b) choices, with a negative peak in the 100–150-ms time window (the N100 component) and a positive peak in the 300–400-ms time window (the P300 component).

When assessing the effect of reward probability on ERPs, we found no univariate differences survived the correction for multiple comparisons in equal ( $p > 0.552$  at all time points, cluster-level permutation test across electrodes and time points) or single-option trials ( $p > 0.175$ , cluster-level permutation test). For equal trials, we found no significant difference in ERPs between preferred and non-preferred



**Fig. 5** Grand-average stimulus-locked ERPs across all EEG electrodes. **a** ERPs from high (100%), medium (80%), and low (20%) reward probability in equal trials. **b** ERPs from high (100%), medium (80%), and low (20%) reward probability in single-option trials. **c** ERPs from equal trials in which the preferred or non-preferred cue was chosen. In all panels, the dashed lines represent standard errors across participants

choices (Fig. 5c,  $p > 0.208$ , cluster-level permutation test). Therefore, in the current study, univariate ERPs were not sensitive to reward probability or preferred/non-preferred choices.

### Multivariate Patterns in Equal Choices

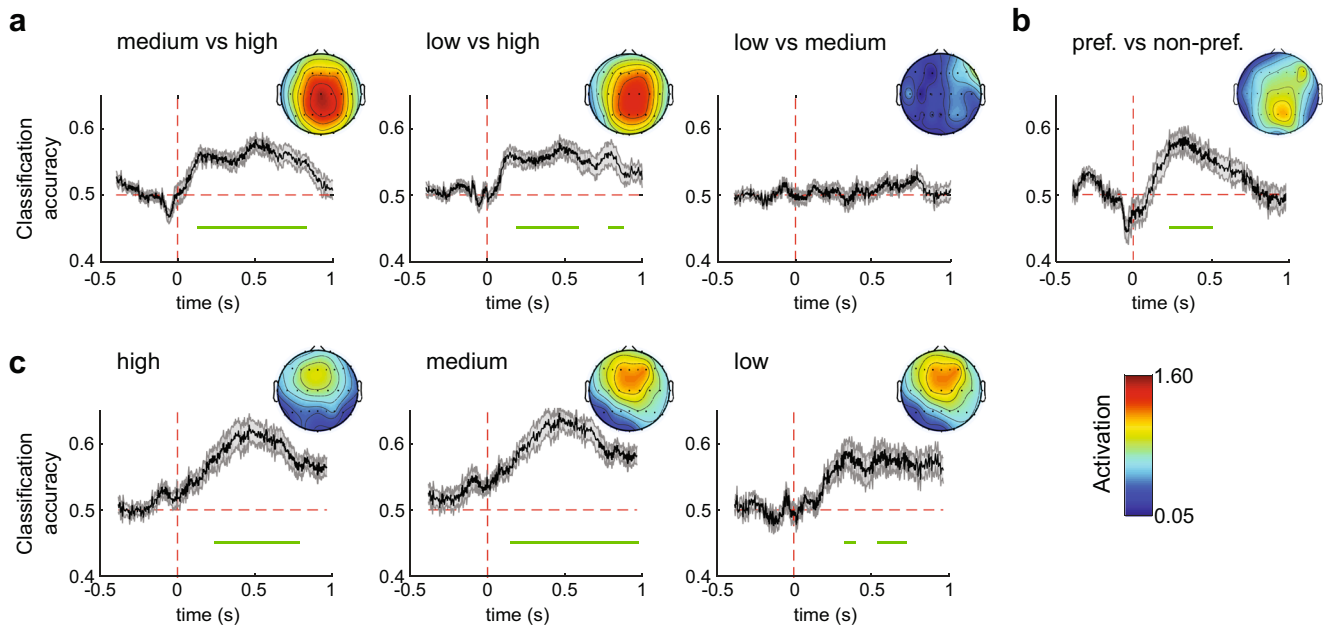
To decode multivariate information representing reward probability in equal choice trials, we applied the linear SVM on multivariate EEG patterns across all electrodes (see “Multivariate Pattern Analysis”). Binary classification between high and medium reward probability was significantly above chance ( $p < 0.01$ , cluster permutation correction, non-parametric Wilcoxon test) from 144 ms after cue onset (Fig. 6a). Similarly, the information between high and low reward probability was decodable above chance from 192 ms after cue onset ( $p < 0.05$ , cluster permutation correction). We found no significant classification accuracy between medium and low reward probability ( $p > 0.16$  in all time points, uncorrected). Therefore, choices associated with certain (100%) rewards were distinguishable from those with uncertain reward probabilities.

We applied a similar classification procedure to decode the information between equal trials in which the participants chose their preferred or non-preferred choices across reward probability levels. The information about preferred versus non-preferred choices was decodable from 316 to 472 ms after cue onset ( $p < 0.009$ , cluster permutation correction).

To evaluate the relative importance of each feature (i.e., EEG electrode) to the classification performance, we calculated the weight vector of SVMs. For each classification problem, we retrained the SVM at each time point with all the data included in the training set and obtained the SVM weight vector. The weight vectors were then transformed into interpretable spatial patterns by multiplying the data covariance matrix (Haufe et al. 2014). The group spatial patterns were calculated by averaging across participants and from all time points which had significant classification accuracy. Relevance spatial patterns based on SVM’s weight vector showed that mid-line central and posterior electrodes contained the most information for significant classification (Fig. 6).

### EEG-Informed Cognitive Modelling

P300 component is a strong candidate for a marker of evidence accumulation. Its amplitude has been associated with attention (Datta et al. 2007), working memory (Kok 2001) and its amplitude with task difficulty (Kok 2001). Prominent models propose it reflects build-to-threshold of the decision variable (Twomey et al. 2015; Kelly and O’Connell 2013) or marks the conclusion of



**Fig. 6** MVPA results. **a** Classification accuracies across time points between equal choices with different levels of reward probability. **b** Classification accuracies across time points between equal trials with preferred and non-preferred choices. **c** Classification accuracies across time points between equal and single-option choices with the same level of reward probability. In all panels, the black lines denote classification accuracies from a stratified 10-fold cross-validation

and the gray areas denote standard errors. Significant decoding time windows (green horizontal bars) were determined from cluster-level permutation tests ( $p < 0.05$ , corrected). Topographic maps represent activation patterns from classification weights, which indicate the contribution of different EEG channels to overall classification accuracies

internal decision-making process (Nieuwenhuis et al. 2005). Considering that the latency of early visual processing is a part of non-decision time (Nunez et al. 2019), we further hypothesized that the evidence accumulation process initiates at N100 peak latency. This led to a theoretical prediction that the slope of the rise in EEG activity between N100 and P300 peak amplitudes reflected the accumulation rate on a trial-by-trial basis. To validate this prediction, we estimated the N100 and P300 components from single trials of equal choices (Fig. 7a), using an SVD-based spatial filter to improve the signal-to-noise ratio of single-trial ERPs (see “Estimation of Single-Trial ERP Components”). This single-trial EEG estimate was then added as a linear regressor (1) of the mean accumulation rate to the LBA model variant with the best fit to behavioral data (i.e., model 15 in Fig. 3a).

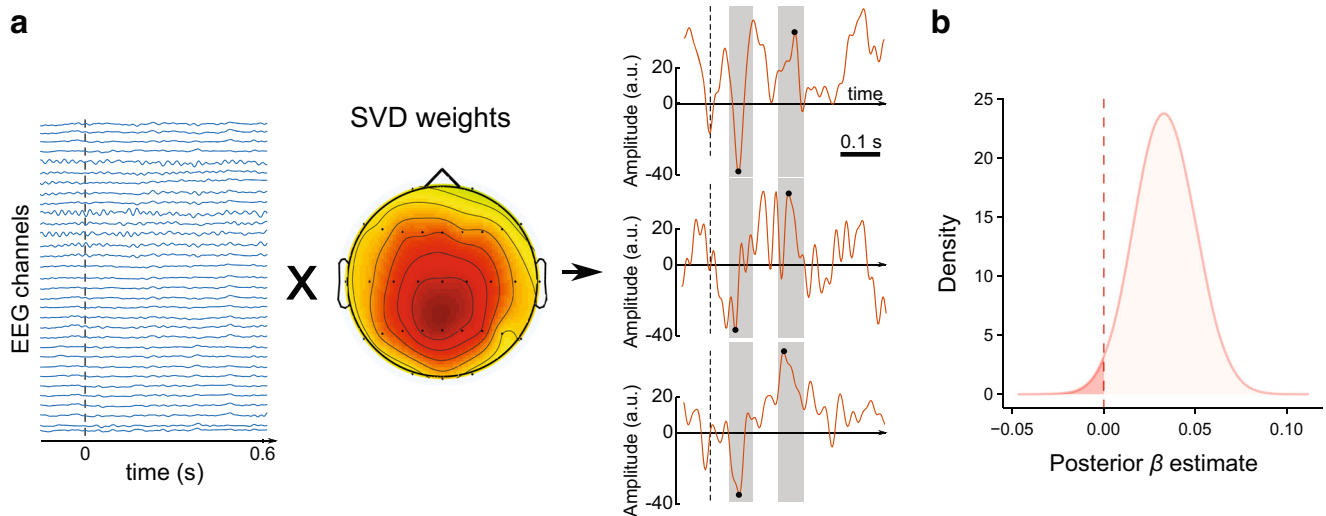
We used the same MCMC procedure to fit the extended LBA model with the EEG-informed regressor to the equal trial data. The extended LBA model showed good convergence ( $\hat{R} \leq 1.02$  for all parameters) and provided a better fit, with a lower LOOIC score 2687 than the model without the EEG-informed regressor (LOOIC score 2796), suggesting that the rising slope of N100–P300 indeed affected the decision process. The posterior estimate of the regression coefficient  $\beta$  provided strong evidence for a positive single-trial effect (Fig. 7b,  $P_{p|D} = 0.983$ ),

indicating that a bigger N100–300 slope is associated with a faster accumulation rate.

### Additional Analyses: Alternative EEG Regressors and Representations of Choice Types

Is it possible that a simpler EEG-based regressor based on a single ERP component could provide a better model fit than the N100–P300 slope? To test this possibility, we fitted four additional extended LBA models with different single-trial EEG regressors applied to the mean accumulation rate: N100 peak latency, N100 peak amplitude, P300 peak latency, and P300 peak latency. All the alternative regression models showed inferior fits (LOOIC scores larger than 2700) than the N100–P300 slope model. We therefore conclude the effects of single-trial EEG activity on the accumulation rate were related to both ERP components.

We did not observe above-chance classification between equal trials with the two levels of uncertain reward probability (Fig. 6a). One may concern whether the lack of significant classification was due to the small number of trials in those conditions. To rule out this possibility, we conducted binary classifications to discriminate equal and single-option trials. The information about trial types (equal vs. single-option) was decodable at every level of reward probability (Fig. 6c,  $p < 0.05$ , cluster corrected),



**Fig. 7** EEG-informed modelling. **a** The schematic diagram of extracting single-trial ERP components. 32-channel EEG signals from a single trial were multiplied by the weights of the first SVD component, calculated from the grand-averaged ERP. Next, the N100 and P300 components in that trial were identified by searching for the peak amplitude in a time of 60–164 ms for the N100 component, and 272–376 ms for the P300 component, respectively. ERP marks

in three representative trials were illustrated in the right column of the panel. The ratio between N100–P300 peak amplitude difference and N100–P300 peak latency difference was calculated as a single-trial regressor for modelling. **b** Posterior estimates of the coefficient between the EEG-informed single-trial regressor (i.e., the rising slope of N100–P300 components) and changes in the accumulation rate

including the one with the least number of trials (i.e., the low reward probability). This result was expected, given the large difference in stimulus presentation and behavioral performance between the two types of choices. SVM-based relevance patterns highlighted the middle central and frontal electrodes to contain most of the information of trial types. These results suggested that the difference in classification accuracies between certain and uncertain reward conditions could not be readily caused by differences in the number of trials.

## Discussion

We provide novel evidence that reward probability and spontaneous preference influence choices between equally probable alternatives and their electrophysiological signatures. We observed two patterns that were consistently distinct at behavioral, cognitive, and neural levels: a certainty effect, distinguishing choices between cues with 100% reward probability and cues with uncertain reward probabilities (80% or 20%), and a preference effect, differentiating between equally valued options. At the behavioral level, reward certainty (i.e., 100% reward vs. non 100% rewards) resulted in disproportionately faster reaction times, while preference biased both choice frequency and RT, resulting in more frequent and faster responses for preferred cues. Using

hierarchical Bayesian implementation of a cognitive model, we showed that reward certainty and preference bias were associated with changes in the accumulation rate, a model-derived parameter to account for the speed of evidence accumulation during decision-making. At the electrophysiological level, the information of certainty and preference could be reliably decoded from multivariate ERP patterns early during decisions, but not from univariate EEG activities. The accumulation rate was further affected by the slope of the rise in ERPs between the N100 and P300 components on a trial-by-trial basis. Together, the current study provides insight into neurocognitive mechanisms driving choices in a deadlock situation, where there is no clear advantage in choosing one option over the other.

The certainty effect implies a monotonic but nonlinear relationship between reward probability and RT in equal choices: the difference between certain (100%) and uncertain (80% and 20%) rewards was greater than that between the two uncertain conditions. This points to a special status of the 100% reward certainty distinct from lower reward probabilities, as the latter always carries a non-zero risk of no reward. The salient representation of the 100% reward certainty is further highlighted by the lack of significant EEG pattern classification between the two uncertain reward probabilities (i.e., 80% vs. 20%, Fig. 6a). Here, the certainty effect in rapid voluntary decisions resembles risk-averse behavior



in economic decisions (Tversky and Kahneman 1989), which overweighs outcomes with 100% certainty relative to probable ones.

Interestingly, reward probability affected RTs across all trial types. It persisted from equal choices to simple reactions to cue locations in single-option trials (Fig. 2d). In unequal choices, there was also a negative association between RT and the sum of reward probability of the two choices (Fig. 2e). Therefore, even though the reward was not contingent upon RT in the current study, we observed a general tendency of accelerating ones' responses in the presence of a more certain reward. These results are akin to the effect of reward magnitude, which also demonstrates a facilitating effect on RT (Schurman and Belcher 1974; Chen and Kwak 2017). In non-human primates, the phasic activation of dopamine neurons in the ventral midbrain has similar response profiles to changes in reward probability and magnitude (Fiorillo et al. 2003), suggesting a common mesolimbic dopaminergic pathway underlying different facets of reward processing that affect decision-making.

Bayesian model comparison identified specific effects of reward probability on accumulation rates, highlighting two possible cognitive origins of the certainty effect. First, in equal choices, cues with 100% certain reward resulted in larger mean accumulation rates than those with uncertain reward probabilities (Fig. 4). Accumulation rate has been linked to the allocation of attention on the task (Schmiedek et al. 2007). Because reward plays a key role in setting both voluntary (top-down) and stimulus-driven (bottom-up) attentional priority (Libera and Chelazzi 2006; Raymond and O'Brien 2009; Krebs et al. 2010; Won and Leber 2016), high reward probability may boost the attentional resources allocated to sensory processing for more rapid decisions. Second, reward probability affected the variability of accumulation rates across trials (Fig. 4). Higher accumulation rate variability has been associated with better-memorized items (Starns and Ratcliff 2014; Osth et al. 2017; Tillman et al. 2017). It is possible that stimuli associated with 100% certain reward were memorized more strongly (Miendlarzewska et al. 2016), a hypothesis to be confirmed in future studies.

Furthermore, MVPA of stimulus-locked ERPs showed multivariate EEG patterns distinguishing between cues with 100% certain reward and other uncertain reward probabilities as early as 150 ms after stimulus onset (Fig 5a, see also Thomas et al. (2013)), and model comparisons found no evidence to support for non-decision time to vary between reward probability levels (Fig. 3a). Considering the average RT of 600 ~ 900 ms in equal choices, our results did not support the latency of post-decision motor preparation, which constitutes a part of the non-decision time (Karahan et al. 2019), to be the source of the certainty effect. This result is consistent with the view that motor

action implementation is independent of the stimulus value (Marshall et al. 2012). Instead, the certainty effect possibly originates from evidence accumulation during the decision process, as supported by the changes in the accumulation rate.

When choosing between equally valued options, classical evidence accumulation theories predict a deadlock scenario with a prolonged decision process (Bogacz et al. 2006). This was not supported by recent experimental findings in value-based decisions (Pirrone et al. 2018; Teodorescu et al. 2016), including the current study, in which equal choices took no longer than unequal ones. Our behavioral, modelling, and EEG analyses indicated a preference bias which could effectively serve as a cognitive mechanism to break the decision deadlock. Compared with non-preferred options, preferred decisions facilitated RTs, were associated with larger accumulation rates, and evoked distinct EEG multivariate patterns. Here, we did not aim to provide a mechanistic interpretation of preference (i.e., why or how the preference bias originated). Instead, our results demonstrated a consistent presence of preference bias before and after cue-probability remapping, independently across reward probabilities (Fig. 2a) and maintained in single-option trials (Fig. 2d), which we considered a novel finding in the literature of voluntary choice.

What can induce a preference bias? Because the cue-probability association was initially randomized and later changed within each session, and no differences in shape preference were found, this bias was not due to stimulus salience but established spontaneously (Voigt et al. 2019). Multiple factors may contribute to the establishment of preferred options. Preference might arise as a function of early choices and outcome frequencies (Izuma et al. 2010; Bakkour et al. 2018), which shape future beliefs or alter the memory trace of certain cue-probability bindings. This interpretation is consistent with an irrationality bias, which favors previously rewarded stimuli, even when controlling for their value (Scholl et al. 2015). Alternatively, some cue-value associations might be remembered more reliably due to a deliberate cognitive strategy of memory resource allocation.

Our results provide little evidence to either support or refute these hypotheses. However, memory strength alone cannot explain the full set of results in the current study. First, it is worth noting that the stimulus-reward mapping was presented a total of 16 times throughout each session (at the beginning of each block and after every 40 trials), and participants took as much time as they needed before the next set of trials. Second, the linear mixed-effect models found significant effects of preference on RT only in equal and single-option trials, but not in unequal trials (Fig. 2). If we were to believe memorization of items to be different between two cues of the same reward

probability, we would expect this to be reflected also in the unequal condition, which was not the case. Future studies could validate these hypotheses by employing more frequent cue-probability remapping throughout experiments and controlling for memory effects. Furthermore, all trials in the current studies were randomized and participants did not have prior knowledge of upcoming stimuli. One future extension would be to evaluate whether presenting prior information of reward probability in an upcoming trial would modulate boundary separation in voluntary decisions, similar to the effect of prior bias on perceptual decisions (Mulder et al. 2012).

The current study considered a simplified form of decision, in which the amount of reward was fixed (i.e., 10 game points). In traditional value-based decisions assumed by the prospect theory, a decision-maker needs to integrate the value and probability of gain or loss to obtain an expected utility for each option (Tversky and Kahneman 1992). Together, our results here and previous studies (Wagner et al. 2020) provide converging evidence that both reward value and probability can influence RT in equal choices. This raises the intriguing possibility of our results to be generalized to choices with the same expected utility but the different combinatorial of value and probability. Interestingly, the multiattribute extension of the LBA model (Trueblood et al. 2014) has been fitted to RTs from such tasks (Cohen et al. 2017), suggesting that our modelling and EEG approaches could also be extended to explore more complex decision problems.

Our study highlights the advantages of EEG-informed cognitive modelling to inform behavioral data. Hierarchical Bayesian parameter estimation of the LBA model provides a robust fit to an individual's behavioral performance with less experimental data needed than other model-fitting methods (Vandekerckhove et al. 2011; Wiecki et al. 2013; Zhang et al. 2016). By integrating single-trial EEG regressors with the cognitive model, we identified the accumulation rate to be affected by the rate of EEG activity changes between visual N100 and P300 components. This result contributes to a growing literature of EEG markers of evidence accumulation processes, including ERP components (Twomey et al. 2015; Loughnane et al. 2016; Nunez et al. 2017), readiness potential (Lui et al. 2018), and oscillatory power (van Vugt et al. 2012). It further consolidates the validity of evidence accumulation as a common computational mechanism leading to voluntary choices of rewarding stimuli (Summerfield and Tsetsos 2012; Afacan-Seref et al. 2018; Maoz et al. 2019), beyond its common applications to perceptually difficult and temporally extended paradigms.

The EEG-informed modelling builds upon the known functional link between the P300 component and evidence

accumulation for decisions (Polich et al. 1996; Verleger et al. 2005; Twomey et al. 2015). A new extension in the current study was to consider the accumulation process begins at the peak latency of the visual N100 component. Theoretically, the delayed initiation of the decision process accounts for information transmission time of 60 ~ 80 ms from the retina (Schmolsky et al. 1998). Single-unit recording concurs with this pre-decision delay, as neurons in putative evidence accumulation regions exhibit a transient dip and recovery activity independent of decisions approximately 90 ms after stimulus onset (Roitman and Shadlen 2002). Practically, our EEG data has a clear N100 component, and time-resolved MVPA identified significant pattern differentiating between task conditions at a similar latency. The relatively early start of the accumulation process in our experiment might be explained by the easily discriminable nature of the cues, consisting of basic shapes with no perceptual noise. Longer visual processing stage has been reported in an experiment involving more complex processing of visual information (Nunez et al. 2019). Further research could dissect the non-decision time (White et al. 2014; Tomassini et al. 2019) and compare latencies of visual encoding across decision tasks with stimuli at different levels of complexity.

Several issues require further consideration. First, our cognitive modeling was not meant to reproduce all the rich behavioral features in the data. To include sufficient observations for model fitting, we combined the data before and after cue-probability remapping. As a result, our model did not account for behavioral changes related to cue remapping. Future studies could employ a multi-session design to investigate how learning new cue-probability associations influences model parameters (Zhang and Rowe 2014).

Second, we focused on the certainty and preference effects by fitting the LBA model only to the data from equal choices. Although simulations indicated that the fitted model provided similar behavioral patterns as in the empirical data in unequal and single-option choices, it was not fitted directly to the experiment data in those two choice conditions. A more parsimonious model for all three types of choices would require additional assumptions, which is beyond the scope of the current study. For example, to incorporate the large RT discrepancy between equal and single-option choices, one could assume that the urgency signal (Boehm et al. 2016; Thura and Cisek 2017) plays a more dominant role in accelerating RT when no apparent comparisons are needed in single-option choices.

Third, our model selection procedure does not encompass all conceivable model types that might account for this data. Independent accumulation is consistent with findings on brain mechanisms of probability-based choices in

humans (Kolling et al. 2014; Scholl et al. 2015), as well as choice behavior in rats (Ojeda et al. 2018). Alternative explanations of certainty and preference effects can be provided by urgency gating (Thura et al. 2012), collapsing threshold (Ratcliff and Frank 2012), or cross-inhibition (Pais et al. 2013; Usher and McClelland 2001). Depending on the parameterization, interpretations based on these models could slightly vary. It is also important to note that there may be no straightforward way to disentangle the interpretations provided by these different models (Miletić and van Maanen 2019). These potential differences however, although important, would not challenge the main conclusions of this paper.

Finally, our model-based analysis is unavoidably constrained by the choice of model, and one needs to be cautious when extending findings to different models. There is an ongoing debate on how accurately different models can mimic each other when estimating the non-decision time (Donkin et al. 2011; Goldfarb et al. 2014; Lerche and Voss 2018). The DDM, for example, tends to predict longer non-decision times than LBA (Dutilh et al. 2019), as well as might be more susceptible to urgency manipulations (Evans 2020). Although an extended DDM has been shown to account for magnitude effects (Ratcliff et al. 2018), the drift rate of a DDM represents the relative signal difference between the two options. As a result, without fitting a new DDM to each condition, the DDM cannot directly describe all conditions in our current study (i.e., the unequal and single-option trials). LBA, on the other hand, assumes an independent accumulator for each option (i.e., each reward cue), offering a parsimonious account to our task and the capacity to produce the qualitative features of responses in all conditions, as demonstrated in our model simulations. It is worth noting that this feature of LBA is shared by other models with multiple accumulators (e.g., the LCA model and the race model). Among those, the LBA model has the simplest form and analytic solutions to the first-passage time, enabling efficient MCMC sampling in Bayesian hierarchical modelling.

In conclusion, when choosing between equally probable reward outcomes, probability and preference selectively modulate the decision processes and their electrophysiological signatures, providing a mechanism for breaking a decision deadlock. These findings extend and substantiate the computational framework of evidence accumulation for voluntary decisions. Our results further highlight the intricate nature of human behavior, as susceptible to external factors as well as endogenous heuristics.

**Supplementary Information** The online version contains supplementary material available at (<https://doi.org/10.1007/s42113-020-00096-6>)

**Acknowledgments** We thank Sabina Baltruschat for assisting in data collection, and Craig Hedge and Petroc Sumner for comments.

**Funding** This study was supported by a European Research Council starting grant (716321). WZ was supported by a PhD studentship from Cardiff University School of Psychology. DK was supported by a PhD studentship from the Engineering and Physical Sciences Research Council (1982622).

**Data Availability** We have made the behavioral and EEG data open access (DOI: 10.6084/m9.figshare.9989552).

**Code Availability** The scripts for behavioral modelling and EEG analysis are open-source and freely available online (<https://github.com/ccbrain/voluntary-decision-eeeg>).

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abler, B., Herrnberger, B., Grön, G., Spitzer, M. (2009). From uncertainty to reward: bold characteristics differentiate signaling pathways. *BMC Neuroscience*, 10(1), 154.
- Afacan-Seref, K., Steinemann, N.A., Blangero, A., Kelly, S.P. (2018). Dynamic interplay of value and sensory information in high-speed decision making. *Current Biology*, 28(5), 795–802.
- Ahn, W.-Y., Haines, N., Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hbayesdm package. *Computational Psychiatry*, 1, 24–57.
- Annis, J., Miller, B.J., Palmeri, T.J. (2017). Bayesian inference with stan: a tutorial on adding custom distributions. *Behavior Research Methods*, 49(3), 863–886.
- Baayen, R.H., Davidson, D.J., Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.
- Bagiella, E., Sloan, R.P., Heitjan, D.F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, 37(1), 13–20.
- Bakkour, A., Zylberberg, A., Shadlen, M.N., Shohamy, D. (2018). Value-based decisions involve sequential sampling from memory. *BioRxiv* page 269290.
- Ballard, I.C., Kim, B., Liatsis, A., Aydogan, G., Cohen, J.D., McClure, S.M. (2017). More is meaningful: the magnitude effect in intertemporal choice depends on self-control. *Psychological Science*, 28(10), 1443–1454.

- Barr, D.J., Levy, R., Scheepers, C., Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bates, D., Kliegl, R., Vasishth, S., Baayen, H. (2015). Parsimonious mixed models. arXiv:1506.04967.
- Bayarri, M.J., & Berger, J.O. (2004). The interplay of Bayesian and frequentist analysis. *Statistical Science*, 19(1), 58–80.
- Betancourt, M. (2017). A conceptual introduction to Hamiltonian monte carlo. arXiv:1701.02434.
- Boehm, U., Hawkins, G.E., Brown, S., van Rijn, H., Wagenmakers, E.-J. (2016). Of monkeys and men: impatience in perceptual decision-making. *Psychonomic Bulletin & Review*, 23(3), 738–749.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700.
- Bogacz, R., Usher, M., Zhang, J., McClelland, J.L. (2007). Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485), 1655–1670.
- Brown, S.D., & Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153–178.
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A. (2017). Stan: a probabilistic programming language. *Journal of Statistical Software*, 76, 1–32.
- Cavanagh, J.F., Wiecki, T.V., Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J. (2011). Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nature Neuroscience*, 14(11), 1462.
- Chen, X.-j., & Kwak, Y. (2017). What makes you go faster?: the effect of reward on speeded action under risk. *Frontiers in Psychology*, 8, 1057.
- Cichy, R.M., Pantazis, D., Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, 17(3), 455.
- Cohen, A.L., Kang, N., Leise, T.L. (2017). Multi-attribute, multi-alternative models of choice: choice, reaction time, and process tracing. *Cognitive Psychology*, 98, 45–72.
- Datta, A., Cusack, R., Hawkins, K., Heutink, J., Rorden, C., Robertson, I.H., Manly, T. (2007). The p300 as a marker of waning attention and error propensity. *Computational Intelligence and Neuroscience*, 2007, 93968.
- Dima, D.C., Perry, G., Messaritaki, E., Zhang, J., Singh, K.D. (2018). Spatiotemporal dynamics in human visual cortex rapidly encode the emotional content of faces. *Human Brain Mapping*, 39(10), 3993–4006.
- Ditterich, J., Mazurek, M.E., Shadlen, M.N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nature Neuroscience*, 6(8), 891–898.
- Donkin, C., Brown, S., Heathcote, A., Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, 18(1), 61–69. <https://doi.org/10.3758/s13423-010-0022-4>.
- Dutilh, G., Annis, J., Brown, S.D., Cassey, P., Evans, N.J., Grasman, R.P.P.P., Hawkins, G.E., Heathcote, A., Holmes, W.R., Kryptos, A.-M., Kupitz, C.N., Leite, F.P., Lerche, V., Lin, Y.-S., Logan, G.D., Palmeri, T.J., Starns, J.J., Trueblood, J.S., van Maanen, L., van Ravenzwaaij, D., Vandekerckhove, J., Visser, I., Voss, A., White, C.N., Wiecki, T.V., Rieskamp, J., Donkin, C. (2019). The quality of response time data inference: a blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*, 26(4), 1051–1069. ISSN 1531-5320. <https://doi.org/10.3758/s13423-017-1417-2>.
- Esber, G.R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B: Biological Sciences*, 278(1718), 2553–2561.
- Evans, N. (2020). Think fast! the implications of emphasizing urgency in decision-making. Psyarchiv. <https://doi.org/10.31234/osf.io/pfrb4>.
- Fiorillo, C.D., Tobler, P.N., Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614), 1898–1902.
- Garrett, D., Peterson, D.A., Anderson, C.W., Thaut, M.H. (2003). Comparison of linear, nonlinear, and feature selection methods for eeg signal classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), 141–144.
- Gelman, A., Rubin, D.B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
- Gold, J.I., & Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5(1), 10–16.
- Gold, J.I., & Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience* 30.
- Goldfarb, S., Leonard, N.E., Simen, P., Caicedo-Núñez, C.H., Holmes, P. (2014). A comparative study of drift diffusion and linear ballistic accumulator models in a reward maximization perceptual choice task. *Frontiers in Neuroscience*, 8, 148–148. ISSN 1662-4548. <https://doi.org/10.3389/fnins.2014.00148>. <https://pubmed.ncbi.nlm.nih.gov/25140124>. 25140124[pmid].
- Gueorguieva, R., & Krystal, J.H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the archives of general psychiatry. *Archives of General Psychiatry*, 61(3), 310–317.
- Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, 87, 96–110.
- Hawkins, G.E., Wagenmakers, E.-J., Ratcliff, R., Brown, S.D. (2015). Discriminating evidence accumulation from urgency signals in speeded decision making. *J Neurophysiol*, 114(1), 40–7.
- Heekeren, H.R., Marrett, S., Ungerleider, L.G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, 9(6), 467–479.
- Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, 107(51), 22014–22019.
- Jahfari, S., Ridderinkhof, K.R., Scholte, H.S. (2013). Spatial frequency information modulates response inhibition and decision-making processes. *PLoS One*, 8(10), e76467.
- Jia, K., Xue, X., Lee, J.-H., Fang, F., Zhang, J., Li, S. (2018). Visual perceptual learning modulates decision network in the human brain: the evidence from psychophysics, modeling, and functional magnetic resonance imaging. *Journal of Vision*, 18(12), 9–9.
- Karahan, E., Costigan, A.G., Graham, K.S., Lawrence, A.D., Zhang, J. (2019). Cognitive and white-matter compartment models reveal selective relations between corticospinal tract microstructure and simple reaction time. *Journal of Neuroscience*, 39(30), 5910–5921.
- Kayser, J., & Tenke, C.E. (2003). Optimizing pca methodology for erp component identification and measurement: theoretical rationale

- and empirical evaluation. *Clinical Neurophysiology*, 114(12), 2307–2325.
- Kelly, S.P., & O'Connell, R.G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience*, 33(50), 19434–19441.
- Kleiner, M., Brainard, D., Pelli, D. (2007). What's new in psychtoolbox-3? *Perception*, 36(14), 1–16.
- Kok, A. (2001). On the utility of p3 amplitude as a measure of processing capacity. *Psychophysiology*, 38(3), 557–577.
- Kolling, N., Wittmann, M., Rushworth, M.F. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*, 81(5), 1190–1202.
- Krajbich, I., Lu, D., Camerer, C., Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3, 193.
- Krebs, R.M., Boehler, C.N., Woldorff, M.G. (2010). The influence of reward associations on conflict processing in the stroop task. *Cognition*, 117(3), 341–347.
- Lerche, V., & Voss, A. (2018). Speed–accuracy manipulations and diffusion modeling: lack of discriminant validity of the manipulation or of the parameter estimates? *Behavior Research Methods*, 50(6), 2568–2585. ISSN 1554-3528. <https://doi.org/10.3758/s13428-018-1034-7>.
- Libera, C.D., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, 17(3), 222–227.
- Loughnane, G.M., Newman, D.P., Bellgrove, M.A., Lalor, E.C., Kelly, S.P., O'Connell, R.G. (2016). Target selection signals influence perceptual decisions by modulating the onset and rate of evidence accumulation. *Current Biology*, 26(4), 496–502.
- Lui, K.K., Nunez, M.D., Cassidy, J.M., Vandekerckhove, J., Cramer, S.C., Srinivasan, R. (2018). Timing of readiness potentials reflect a decision-making process in the human brain. bioRxiv page 338806.
- Mangun, G.R., & Hillyard, S.A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human Perception and Performance*, 17(4), 1057.
- Maoz, U., Yaffe, G., Koch, C., Mudrik, L. (2019). Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice. *eLife* 8.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190.
- Marshall, J.A., Bogacz, R., Gilchrist, I.D. (2012). Consistent implementation of decisions in the brain. *PLoS One*, 7(9), e43443.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, 94, 305–315.
- Miendlarzewska, E.A., Bavelier, D., Schwartz, S. (2016). Influence of reward motivation on human declarative memory. *Neuroscience & Biobehavioral Reviews*, 61, 156–176.
- Miletić, S., & van Maanen, L. (2019). Caution in decision-making under time pressure is mediated by timing ability. *Cognitive Psychology*, 110, 16–29. ISSN 0010-0285(Print). <https://doi.org/10.1016/j.cogpsych.2019.01.002>.
- Mulder, M.J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., Forstmann, B.U. (2012). Bias in the brain: a diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32(7), 2335–2343.
- Nieuwenhuis, S., Aston-Jones, G., Cohen, J.D. (2005). Decision making, the p3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, 131(4), 510.
- 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.
- Nunez, M.D., Gosai, A., Vandekerckhove, J., Srinivasan, R. (2019). The latency of a visual evoked potential tracks the onset of decision making. *Neuroimage*, 197, 93–108.
- Ojeda, A., Murphy, R.A., Kacelnik, A. (2018). Paradoxical choice in rats: subjective valuation and mechanism of choice. *Behavioural Processes*, 152, 73–80.
- Osth, A.F., Dennis, S., Heathcote, A. (2017). Likelihood ratio sequential sampling models of recognition memory. *Cognitive Psychology*, 92, 101–126.
- Oud, B., Krajbich, I., Miller, K., Cheong, J.H., Botvinick, M., Fehr, E. (2016). Irrational time allocation in decision-making. *Proceedings of the Royal Society B: Biological Sciences*, 283(1822), 20151439.
- Pais, D., Hogan, P.M., Schlegel, T., Franks, N.R., Leonard, N.E., Marshall, J.A. (2013). A mechanism for value-sensitive decision-making. *PLoS One*, 8(9), e73216.
- Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P. (2005). Recipes for the linear analysis of EEG. *Neuroimage*, 28(2), 326–341.
- Phillips, H., Cope, T., Hughes, L., Zhang, J., Rowe, J. (2018). Monitoring the past and choosing the future: the prefrontal cortical influences on voluntary action. *Scientific Reports*, 8(1), 1–13.
- Pirrou, A., Azab, H., Hayden, B.Y., Stafford, T., Marshall, J.A. (2018). Evidence for the speed–value trade-off: human and monkey decision making is magnitude sensitive. *Decision*, 5(2), 129.
- Pisauro, M.A., Fouragnan, E., Retzler, C., Philiastides, M.G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nature Communications*, 8(1), 1–9.
- Polanía, R., Krajbich, I., Grueschow, M., Ruff, C.C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, 82(3), 709–720.
- Polich, J., Ellerson, P.C., Cohen, J. (1996). P300, stimulus intensity, modality, and probability. *International Journal of Psychophysiology*, 23(1–2), 55–62.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59.
- Ratcliff, R., & Frank, M.J. (2012). Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neuro-computational and diffusion models. *Neural Computation*, 24(5), 1186–1229.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922.
- Ratcliff, R., & Smith, P.L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111(2), 333.
- Ratcliff, R., Voskuilen, C., Teodorescu, A. (2018). Modeling 2-alternative forced-choice tasks: accounting for both magnitude and difference effects. *Cognitive Psychology*, 103, 1–22.
- Raymond, J.E., & O'Brien, J.L. (2009). Selective visual attention and motivation: the consequences of value learning in an attentional blink task. *Psychological Science*, 20(8), 981–988.
- Roitman, J.D., & Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience*, 22(21), 9475–9489.
- Rustichini, A. (2009). Neuroeconomics: formal models of decision making and cognitive neuroscience. In *Neuroeconomics* (pp. 33–46). Elsevier.
- Schiebener, J., & Brand, M. (2015). Decision making under objective risk conditions—a review of cognitive and emotional correlates, strategies, feedback processing, and external influences. *Neuropsychology Review*, 25(2), 171–198.

- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., Wittmann, W.W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, 136(3), 414.
- Schmolecky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., Leventhal, A.G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, 79(6), 3272–3278.
- Scholl, J., Kolling, N., Nelissen, N., Wittmann, M.K., Harmer, C.J., Rushworth, M.F. (2015). The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. *Journal of Neuroscience*, 35(32), 11233–11251.
- Schurman, D., & Belcher, J. (1974). Noncontingent reward magnitude effects on reaction time: a replication and extension. *Bulletin of the Psychonomic Society*, 3(2), 104–106.
- Shadlen, M.N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90(5), 927–939.
- Starns, J.J., & Ratcliff, R. (2014). Validating the unequal-variance assumption in recognition memory using response time distributions instead of ROC functions: a diffusion model analysis. *Journal of Memory and Language*, 70, 36–52.
- Summerfield, C., & Tsetsos, K. (2012). Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Frontiers in Neuroscience*, 6, 70.
- Teodorescu, A.R., Moran, R., Usher, M. (2016). Absolutely relative or relatively absolute: violations of value invariance in human decision making. *Psychonomic Bulletin & Review*, 23(1), 22–38.
- Thaler, R.H. (1991). Some empirical evidence on dynamic inconsistency. *Quasi Rational Economics*, 1, 127–136.
- Thomas, J., Vanni-Mercier, G., Dreher, J.-C. (2013). Neural dynamics of reward probability coding: a magnetoencephalographic study in humans. *Frontiers in Neuroscience*, 7, 214.
- Thura, D., & Cisek, P. (2017). The basal ganglia do not select reach targets but control the urgency of commitment. *Neuron*, 95(5), 1160–1170.
- Thura, D., Beauregard-Racine, J., Fradet, C.-W., Cisek, P. (2012). Decision making by urgency gating: theory and experimental support. *Journal of Neurophysiology*, 108(11), 2912–2930.
- Tillman, G., Osth, A.F., van Ravenzwaaij, D., Heathcote, A. (2017). A diffusion decision model analysis of evidence variability in the lexical decision task. *Psychonomic Bulletin & Review*, 24(6), 1949–1956.
- Tobler, P.N., O’Doherty, J.P., Dolan, R.J., Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, 97(2), 1621–1632.
- Tomassini, A., Price, D., Zhang, J., Rowe, J.B. (2019). On the evolution of neural decisions from uncertain visual input to uncertain actions. bioRxiv page 803049.
- Trueblood, J.S., Brown, S.D., Heathcote, A. (2014). The multi-tribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, 121(2), 179.
- Tversky, A., & Kahneman, D. (1989). Rational choice and the framing of decisions. In *Multiple criteria decision making and risk analysis using microcomputers* (pp. 81–126). Springer.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- 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.
- Usher, M., & McClelland, J.L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, 108(3), 550.
- van Inwagen, P. (1989). When is the will free? *Philosophical Perspectives*, 3, 399–422. ISSN 15208583, 17582245. <http://www.jstor.org/stable/2214275>.
- van Vugt, M.K., Simen, P., Nystrom, L.E., Holmes, P., Cohen, J.D. (2012). Eeg oscillations reveal neural correlates of evidence accumulation. *Frontiers in Neuroscience*, 6, 106.
- Vandekerckhove, J., Tuerlinckx, F., Lee, M.D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, 16(1), 44.
- Vehtari, A., Gelman, A., Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5), 1413–1432.
- Verleger, R., Jaśkowski, P., Wascher, E. (2005). Evidence for an integrative role of p3b in linking reaction to perception. *Journal of Psychophysiology*, 19(3), 165–181.
- Voigt, K., Murawski, C., Speer, S., Bode, S. (2019). Hard decisions shape the neural coding of preferences. *Journal of Neuroscience*, 39(4), 718–726.
- Wagner, B.J., Clos, M., Sommer, T., Peters, J. (2020). Dopaminergic modulation of human inter-temporal choice: a diffusion model analysis using the d2-receptor-antagonist haloperidol. bioRxiv.
- Wheeler, D.R. (1974). Brand loyalties: qualitative, quantitative, or both? *Journal of the Academy of Marketing Science*, 2(4), 651–658.
- White, C.N., Congdon, E., Mumford, J.A., Karlsgodt, K.H., Sabb, F.W., Freimer, N.B., London, E.D., Cannon, T.D., Bilder, R.M., Poldrack, R.A. (2014). Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. *Journal of Cognitive Neuroscience*, 26(8), 1601–1614.
- Wiecki, T.V., Sofer, I., Frank, M.J. (2013). HDDM: hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics*, 7, 14.
- Won, B.-Y., & Leber, A.B. (2016). How do magnitude and frequency of monetary reward guide visual search? *Attention, Perception, & Psychophysics*, 78(5), 1221–1231.
- Young, M.E., Webb, T.L., Rung, J.M., McCoy, A.W. (2014). Outcome probability versus magnitude: when waiting benefits one at the cost of the other. *PloS One*, 9(6), e98996.
- Zhang, J. (2012). The effects of evidence bounds on decision-making: theoretical and empirical developments. *Frontiers in Psychology*, 3, 263.
- Zhang, J., & Rowe, J.B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in Neuroscience*, 8, 69.
- Zhang, J., & Rowe, J.B. (2015). The neural signature of information regularity in temporally extended event sequences. *Neuroimage*, 107, 266–276.
- Zhang, H., & Wang, Z. (2011). A normal distribution-based over-sampling approach to imbalanced data classification. In *International conference on advanced data mining and applications* (pp. 83–96). Springer.
- Zhang, J., Rittman, T., Nombela, C., Fois, A., Coyle-Gilchrist, I., Barker, R.A., Hughes, L.E., Rowe, J.B. (2016). Different decision deficits impair response inhibition in progressive supranuclear palsy and Parkinson’s disease. *Brain*, 139(1), 161–173.