Tracing the neural trajectories of evidence accumulation and motor preparation processes during voluntary decisions

Lauren C. Fong[1*], Paul M. Garrett[1], Philip L. Smith[1], Robert Hester[1], Stefan Bode[2], and Daniel Feuerriegel[1]

[1] Melbourne School of Psychological Sciences, The University of Melbourne, Australia

The University of Melbourne, 17 Spencer Road, Parkville VIC 3010, Australia

[2] Division of Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Saadiyat Marina District - Abu Dhabi - United Arab Emirates

[*]Correspondence: fonglc@unimelb.edu.au

Abstract

Voluntary decisions have previously been described by *where* they arise in the brain and *how* actions corresponding to one's choice are prepared. However, the processes by which these internally guided decisions are formed and translated into action remain unclear. We tested whether neural signatures of evidence accumulation and motor preparation – established in perceptual decision-making – are also observed during voluntary decisions. Forty-nine adults made voluntary (free choice between two options) and forced (single specified choice) decisions while electroencephalography was recorded. To isolate decision formation from action selection, all responses were made with a right-handed keypress. After applying signal deconvolution and current source density transformations, we examined three signals: the centro-parietal positivity (CPP), Mu/Beta (8 – 30 Hz) amplitudes, and a left-hemisphere readiness potential contralateral to the response hand. For both voluntary and forced decisions, CPP and Mu/Beta amplitudes showed hallmark accumulation-to-bound dynamics – faster responses had steeper ramping signals that converged to stereotyped pre-response amplitude levels. The left-hemisphere readiness potential likewise showed convergence to a pre-response level but with less reliable evidence for RT-dependent slope scaling, resembling a late, effector-specific motor gate rather than an evolving decision variable. Waveform amplitudes across the time course of these signals also did not differ reliably between voluntary and forced decisions. Our results indicate that, like perceptual decisions, voluntary choices are also formed by a graded accumulation process that culminates in a motor action, with the CPP indexing an evolving decision variable and Mu/Beta indexing effector-specific motor readiness.

Keywords: CPP, RP, Mu/Beta, Evidence Accumulation, Motor Preparation, Voluntary Decisions

1 Introduction

Every day, we make countless voluntary decisions. Some are trivial – like choosing which mug to drink coffee from – and others more consequential – such as whether to speak publicly about a societal issue or when to tackle a difficult task. Voluntary decisions require us to first recognise that more than one course of action exists – two coffee mugs – then evaluate and deliberate between options by gathering evidence towards each alternative. This process is guided by internally generated goals and preferences (e.g., I prefer the blue mug over the green mug, or I need the insulated mug to keep my coffee warm) and to a lesser extent, by situational context and external cues (Haggard, 2008). Once sufficient evidence has been accumulated to reach a decision, we commit to a choice by executing the corresponding motor action (Bode et al., 2014; Bogler et al., 2023; Brass et al., 2019; Roskies, 2010; Schurger et al., 2012). In this article, we detail the neural dynamics underlying decisional evidence accumulation in voluntary decision making.

Research on the neurophysiology of volition typically contrasts voluntary (i.e., free-choice) trials with externally cued control trials to isolate the neural processes underlying internally guided decisions. In voluntary trials, one option is selected from among many; in control trials, only one option is available. Functional MRI studies consistently show that voluntary trials engage mesial frontal regions – including the pre-supplementary motor area (pre-SMA), supplementary motor area (SMA), and anterior cingulate cortex – and parietal areas (see Seghezzi et al., 2024). Multivariate decoding analyses have further demonstrated that choice-selective information can be detected in these areas prior to action, predicting object choices (Bode et al., 2013; Löffler et al., 2020) and arithmetic operations (Haynes et al., 2007; Soon et al., 2013).

Work using electroencephalography (EEG) further provides millisecond-level resolution of motor preparation preceding voluntary actions. One key event-related potential (ERP) component is the readiness potential (RP). This is a slow, negative-going

wave measured over central electrodes (e.g., Cz) that gradually increases in amplitude prior to movement (Deecke et al., 1969; Kornhuber & Deecke, 1965) and is thought to originate in the pre-SMA/SMA (Cui et al., 1999; Fried et al., 2011). The RP is larger in amplitude for spontaneous compared to instructed actions (e.g., Dirnberger et al., 2008; Jahanshahi et al., 1995), deliberate versus arbitrary actions (e.g., Masaki et al., 1998; Travers et al., 2021), and when decisions were made with less reliance on exogenous information (e.g., Gluth et al., 2013; Travers & Haggard, 2021). Consequently, the RP has been widely interpreted as an index of motor preparation for voluntary actions.

Despite clear evidence of where voluntary actions are implemented in the brain, the processes of decision formation that precede voluntary action remain underspecified. Classic EEG studies have emphasised action execution markers, like the RP, without characterising the evolving decision dynamics that precede motor implementation (Steinemann et al., 2018), particularly when stimulus-response mappings or effector demands vary across conditions. This is in direct contrast to perceptual forced-choice decision-making studies that have identified neural signals that reflect evidence accumulation dynamics (Kelly et al., 2021), as formalised in sequential sampling models such as the Diffusion Decision Model (DDM; Ratcliff, 1978; Ratcliff and Smith, 2004).

Sequential sampling models jointly explain choice and response times (RTs) using psychologically interpretable parameters: decisions are reached after noisy evidence accumulates over time and crosses a decision-boundary, triggering a motor response (Ratcliff et al., 2016; Smith & Ratcliff, 2004). Hypothesised neural correlates of evidence accumulation trajectories exhibit two defining properties: (i) build-up rates that scale inversely with RTs, often mapping onto the evidence quality (drift rate) used to inform the decision (O'Connell & Kelly, 2021; O'Connell et al., 2012), and (ii) pre-response amplitudes that converge at a common level, consistent with a fixed decision boundary parameter (Kelly & O'Connell, 2013; Philiastides et al., 2014).

The centro-parietal positivity (CPP) has been identified as a neural correlate of an evolving decision variable that exhibits both of these properties. The CPP is a positive-going ERP component measured over midline posterior electrodes (e.g., Pz) that ramps towards stereotyped amplitude levels prior to a response, regardless of RT (Kelly & O'Connell, 2013; O'Connell et al., 2012). This ramping is steeper for faster than for slower decisions (O'Connell & Kelly, 2021; O'Connell et al., 2012), consistent with how higher rates of evidence accumulation produce faster responses. CPP dynamics are dissociable from both sensory encoding and motor preparatory activity, making the CPP a relatively direct, supramodal neural correlate of decision formation (O'Connell & Kelly, 2021; Steinemann et al., 2018). While the CPP has been robustly established in perceptual tasks (e.g., Feuerriegel et al., 2021; Kelly & O'Connell, 2013; Kelly et al., 2021; Steinemann et al., 2018; Twomey et al., 2015), its generalisation beyond perceptual decision-making remains under active debate (Frömer et al., 2024). Despite being previously reported in value-based (Pisauro et al., 2017) and memory-based decisions (van Ede & Nobre, 2024; van Vugt et al., 2019), methodological considerations have prompted caution about interpreting the CPP as a domain-general evidence accumulation decision variable (for further discussion, see O'Connell et al., 2025).

Mu/Beta (MB; 8–30 Hz) spectral amplitudes also reflect accumulation-to-bound dynamics in decision-making tasks where motor responses are required. MB amplitudes gradually decrease prior to an action at electrodes contralateral to the response hand (e.g., C3 and C4), indexing the progressive release of motor inhibition (Donner et al., 2009; Lange et al., 2013; Pfurtscheller & Lopes da Silva, 1999). Like the CPP, the rate of amplitude change scales with RT and reaches a fixed amplitude prior to a motor response (Donner et al., 2009; Lange et al., 2013; O'Connell et al., 2012). Contrary to the CPP, the dynamics of MB activity are sensitive to time constraints and urgency (Kelly et al., 2021; Murphy et al., 2016; Steinemann et al., 2018), consistent with adaptive changes in motor readiness that are responsive to different task demands. Collectively, these findings support

MB amplitudes as a motor preparatory signal that accumulates towards a motor-execution boundary and adapts to task dynamics.

Beyond its classic descriptor as a marker of voluntary actions, the RP has been conceptualised within an evidence accumulation framework for simple self-paced actions, modelled as a leaky stochastic accumulator (Schurger, 2018; Schurger et al., 2012). Based on this view, stochastic moment-to-moment fluctuations of neural noise in motor circuits act as "evidence" that ramps toward a motor execution boundary. Once these fluctuations happen to cross the boundary, an action is initiated – analogous to motor threshold-crossing dynamics shown in MB amplitudes described earlier. Under ambiguous or absent external evidence, or in arbitrary decisions with no strong reasons to pick one option over another, threshold-crossing would be primarily noise-driven, and facilitated by lowered boundaries (Brass et al., 2019; Maoz et al., 2019).

This threshold-crossing notion has been extended from arbitrary, spontaneous voluntary actions to more meaningful voluntary decisions. The conditional intention and integration to bound (COINTOB) model proposed by Brass et al. (2019) conceptualises the RP as a neural signature of a general accumulation-to-bound process that is dependent on evidence inputs and policies. In arbitrary, self-paced decision contexts, the accumulator would be dominated by stochastic neural fluctuations; in deliberate decisions, the same evidence accumulation mechanism integrates decision-relevant endogenous (e.g., preferences or goals) and/or external evidence (e.g., monetary values or visual choice stimuli), alongside some noise. The COINTOB model also accommodates lowered or collapsing bounds in decisions made under speed pressure – dynamics similarly observed in MB amplitudes under urgency manipulations. Thus, differences across decision contexts may reflect changes in inputs and bound policies, yet underlie similar evidence accumulation mechanisms.

Empirically, this notion has been supported by recent work showing that single-trial

RP durations predict model-derived decision times in voluntary actions (Lui et al., 2021). Racing-accumulator model fits have shown that arbitrary decisions are well captured by noise-driven accumulation, but less so for deliberate decisions, which were better captured by a value-assessment component (Maoz et al., 2019). A lateralised variant of the RP (LRP), derived from the contralateral-ipsilateral difference over motor electrodes (Eimer, 1998; Haggard, 2009), likewise shows accumulation-to-bound dynamics that can scale with sensory evidence in motion discrimination tasks and lags behind the CPP (Kelly & O'Connell, 2013). The motor system is also thought to receive and track evolving decision evidence during deliberation to prime motor preparation and reduce computational costs when an action is required (Selen et al., 2012). Together, these observations make it plausible that RP/LRP can, in some contexts, index decision formation within the motor system, complementing the CPP and MB signals.

Taken together, the CPP, MB, and L/RP have been established as neural correlates of decision evidence accumulation (CPP) and motor execution dynamics (MB, L/RP) in perceptual decision-making. However, unlike in perceptual decisions, voluntary decisions are primarily endogenously driven, lack objectively correct outcomes, and often unfold without strict deadlines or explicit speed-accuracy constraints. Although these distinctions do not necessarily imply differences in neural correlates of evidence accumulation and motor execution between perceptual and voluntary decisions, they raise a critical question: do these signals extend to voluntary decisions and, if so, do they also exhibit canonical accumulation-to-bound properties?

We address these open questions by recording EEG from a large sample of participants (N=49) who completed a colour decision task featuring both voluntary (two options) and forced (one option) trials. A right-handed keypress was always used to denote when a decision was made, allowing us to focus on a left-hemisphere readiness potential (LHRP) as a contralateral motor-preparatory measure while equating action selection and

stimulus-response mapping. This also allowed us to assess the dynamics of voluntary decisions that did not co-occur with action selection. We then tested whether the three candidate signals – CPP, MB, and LHRP – exhibited the two canonical characteristics of an evolving decision variable in voluntary decisions: (a) build-up rates that scale inversely with RTs, and (b) pre-response amplitudes that build to a fixed level, irrespective of RTs (Kelly & O'Connell, 2013). To disentangle EEG signals time-locked to stimulus onset that may overlap with non-stimulus-locked activity (addressing concerns in Frömer et al., 2024), we used Residue Iteration Decomposition (RIDE) to deconvolve different contributions to EEG waveforms (Ouyang et al., 2015). We also applied current source density (CSD) transformations to enhance spatial precision and reduce volume conduction. This allowed us to characterise whether these candidate signals trace the trajectories of an evolving decision variable in endogenously driven voluntary decisions under tightly controlled motor demands. We could further compare these dynamics to forced decisions in which there is a single, externally driven choice.

2 Method

2.1 Participants

We recruited 50 participants (17 men and 33 women) aged between 18-39 years (M_{age} = 20.88). All participants had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders; 42 participants were right-handed. We excluded one participant due to excessively noisy EEG data, leaving 49 participants for analyses. Each participant received either $30 AUD or two course credits. The University of Melbourne Human Research Ethics Committee approved this study (Approval Number 27017).

2.2 Stimuli and Apparatus

We presented single- and dual-coloured balloons against a black background (Figure 1). Four colours - blue, green, orange, and pink - were chosen from a CIELAB colour wheel (Wyszecki & Stiles, 1982) sampled 90° apart, at 43° (orange), 133° (green), 223° (blue) and 313° (pink). They were then converted to RGB values: orange = [206, 104, 0], green = [19, 144, 0], blue = [0, 176, 238], and pink = [236, 0, 216]. The wheel was centred at L = -54, a = 18, b = -8 with a radius of 108.82 mm (as used in Adam et al., 2017), to ensure approximately equal luminance and distinct colours to reduce saliency biases in participants' choices. Stimuli were presented using Psychtoolbox (Brainard, 1997) interfacing MATLAB R2022b (Mathworks Inc.) on a 24.5-inch ASUS ROG Swift PG258Q monitor (1920 x 1080 pixels, 60 Hz refresh rate). Participants responded using a Keychron Q0 QMK Custom Number Pad (1000 Hz polling rate).

2.3 Procedure

Participants were seated 50 cm from the monitor in a dark, soundproof room, with their heads stabilised using a chin rest. The experiment included two decision types: voluntary decisions and forced decisions. In voluntary trials (Figure 1A), participants viewed a dual-coloured balloon, indicating two available colour choices. In forced trials, they saw only a single colour (Figure 1B). To encourage more deliberate choice behaviour while keeping voluntary decisions largely free from external constraints (Haggard, 2008), participants were given a cover story to "*collect a colourful range of balloons to decorate a rainbow-themed party*". This also avoided explicit value-based elements (e.g., rewards or penalties) that could bias strategic or reward-related choice behaviour.

Each trial began with a white fixation stimulus presented for 1200 ms. We instructed participants to maintain a central fixation whenever it appeared. The decision phase then began with colour choices presented on-screen until participants pressed the

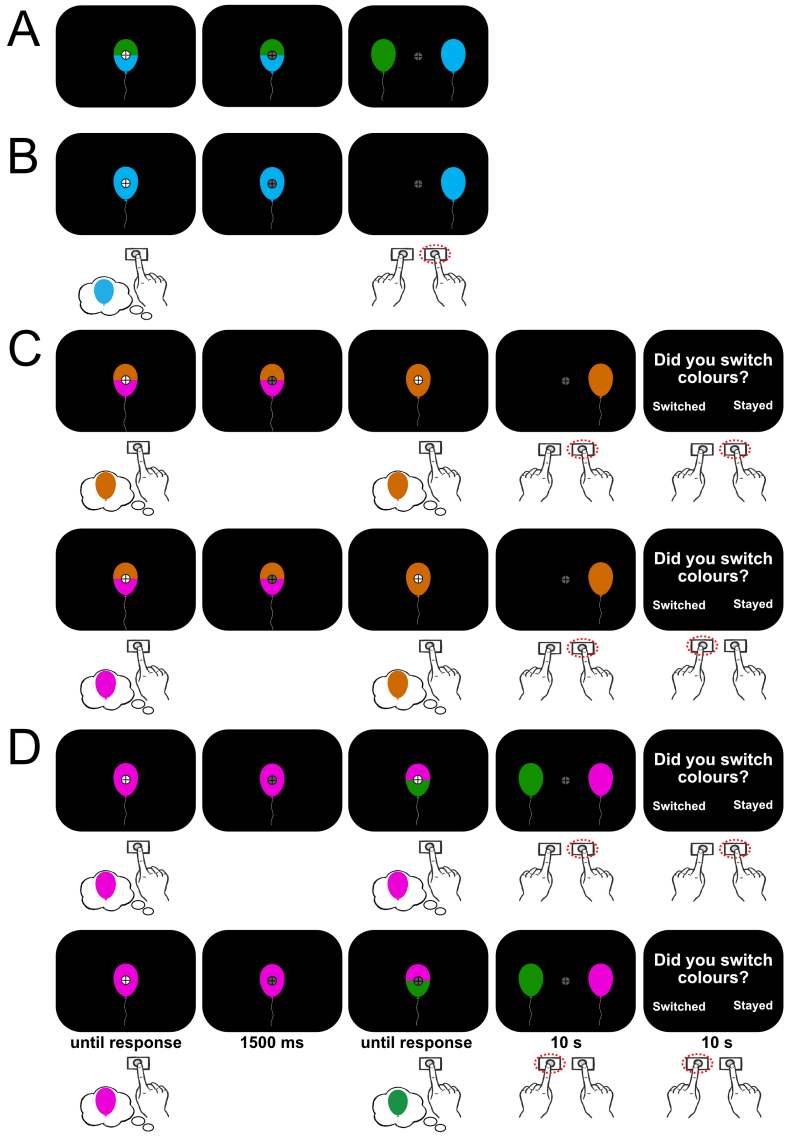


Figure 1

Task design. A) Voluntary decision trials. B) Forced decision trials. In each trial, participants were shown a dual- (voluntary) or single-coloured (forced) balloon and made a keypress with their right middle finger as soon as a decision was made. Balloons of a single colour were then presented at the left and/or right of fixation and participants then indicated the colour they had originally chosen. C) Voluntary-to-forced decision trials. After an initial voluntary decision was made, one of the two colours was removed, forcing participants to either maintain their initial choice or switch. D) Forced-to-voluntary decision trials. After the sole presented option was chosen, a new colour was introduced, allowing participants to maintain or switch their initial choice. In these trials, participants indicated their final choice when balloons appeared to the left and/or right of fixation. They subsequently indicated whether their final choice was the same or different to their initial choice.

“+” key on the number pad with their right middle finger to record the moment a decision had been made. Deadlines were not imposed during this decision phase to preserve the voluntary nature of the decisions and to avoid introducing urgency effects. Responses were constrained to the right hand to control for stimulus-response mapping and action selection. Following their response, the fixation stimulus turned grey for 1500 ms and the balloon remained on-screen. Following this, either one or two balloons (corresponding to the colour choices in the trial) were presented to the left and/or right of fixation. Participants pressed keys with their left or right index fingers to indicate the colour they had originally chosen in that trial.

On half of the trials, a second "change-of-mind" (CoM) decision phase followed. For initially voluntary trials, one of the two initial options was removed, forcing participants to either maintain their original choice or switch to the remaining option (Figure 1C). For initially forced trials, a new colour option was introduced, allowing participants to either maintain or switch their initial choice (Figure 1D). Following this CoM decision phase, we presented the available option(s) to the left and right of fixation, and participants indicated their chosen colour using left- or right-labelled keys (i.e., "1" and "3" keys) with their corresponding index fingers. Participants then reported whether they had switched or stayed with their initial decisions. Participants had 10 seconds to report their colour choices and whether they had switched colour choices.

Participants completed 8 blocks of 36 trials (288 trials in total). Each block lasted approximately 6 minutes, followed by a self-paced break (minimum 10 seconds). Before the main experiment, participants completed one practice block of 36 trials, with the option to attempt up to two additional practice blocks. We report analyses of EEG data from the initial decision phase only. Analyses of EEG data from the CoM decision phase were not directly relevant to our research questions and are not reported here but will be the focus of a future publication.

2.4 EEG Data Acquisition and Processing

We recorded EEG at a sampling rate of 512 Hz using a BioSemi ActiveTwo system with 64 scalp electrodes, using common mode sense (CMS) and driven right leg (DRL) electrodes (http://biosemi.com/faq/cms/&/drl.htm). We placed an additional eight electrodes on the left and right mastoids, 1 cm from the outer canthi of each eye, and above and below the centre of each eye.

We processed the data using EEGLAB v2023.1 (Delorme & Makeig, 2004) running in Matlab R2022b. First, we identified excessively noisy channels by visual inspection (mean number of bad channels = 0.24, range 0–3) and excluded these from average reference calculations and Independent Component Analysis (ICA). Next, we removed sections with large amplitude artefacts by visual inspection, referenced the data to the average of all the channels, and low-pass filtered the data at 40 Hz. We removed channel AFz to compensate for the rank deficiency introduced by average referencing. We then duplicated the dataset for ICA and high-pass filtered at 0.1 Hz to improve stationarity for ICA. We identified and removed independent components associated with eye blinks and saccades from the unfiltered dataset, high-pass filtered the data at 0.1 Hz, and interpolated AFz and any noisy channels using spherical spline interpolation. With the resulting data, we created stimulus-locked segments of EEG data spanning –2000 to +4000 ms relative to stimulus onset in the initial decision phase and baseline corrected using the 200 ms pre-stimulus interval. Epochs with signals at any scalp channel exceeding ±200 µV were excluded.

In speeded decision-making tasks, EEG signals time-locked to stimulus onset can influence ERP waveforms that are time-locked to the decision and motor response. To minimise the influence of overlapping stimulus-locked activity, we used the RIDE toolbox (Ouyang et al., 2015), following the approach of Steinemann et al. (2018). The RIDE toolbox employs an algorithm that deconvolves stimulus-locked and non-stimulus-locked

EEG signals, and has superior performance over comparable methods (Ouyang et al., 2017, 2020). We used the RIDE algorithm to decompose single-trial stimulus-locked data into three subcomponents for each participant using the following time windows: (1) an S (stimulus-locked) subcomponent (0 to 800 ms post-stimulus), (2) an R (response-locked) subcomponent (-600 to -200) ms relative to response onset), and (3) a C (time-varying) subcomponent (200 to 1000 ms post-stimulus) that captures neural activity that differs in latency from trial to trial (e.g., evolving decision formation processes) and not strictly time-locked to either stimulus or response. We included the three subcomponents to improve the accuracy of the S subcomponent estimate in the presence of co-occurring, non stimulus-locked activity that is captured by the C and R subcomponents. Following Sun et al. (2024), we used an extended estimation window for the S subcomponent, relative to default window recommendations in Ouyang et al. (2015), to conservatively estimate and remove stimulus-aligned EEG signals. This addresses concerns noted by Frömer et al. (2024) that response-aligned signals (e.g., evidence accumulation) may artefactually reflect overlapping stimulus-evoked activity. Following best practices implemented in Steinemann et al. (2018) and Sun et al. (2024), we subtracted the estimated S subcomponent (see Supplementary Figure S1A and B) from each stimulus-locked epoch before deriving stimulus- and response-locked segments of EEG data. This subtraction removes estimated stimulus-locked EEG signals, allowing the remaining waveforms to more selectively reflect time-varying and response-locked dynamics (i.e., decision formation and motor preparation). Note that we did not use the R or C subcomponent estimates for our analyses. We also provide plots without ERP deconvolution applied for interested readers in Supplementary Figure S1C–H.

Finally, we created response-locked epochs spanning -1200 to -300 ms relative to the response from the resulting S-subcomponent-subtracted stimulus-locked epochs to retain the same pre-stimulus baseline for both stimulus- and response-locked signals (as done in Feuerriegel et al., 2022). This approach allowed us to use a baseline for response-locked

epochs that minimises contamination from ongoing motor processes – especially pertinent for voluntary decisions, where the onset of volition is difficult to determine (Haggard, 2008). We then converted the epoched data into current source density (CSD) estimates using the CSD Toolbox (Kayser & Tenke, 2006) with an m-constant = 4 and $\lambda = 0.00001$. CSD highlights local voltage changes and diminishes the effect of diffuse activity, thereby improving the topographical resolution of EEG signals. This step was crucial for minimising signal overlap between co-occurring signals generated at frontal and central electrodes (e.g., RP and MB Power) and more posterior electrodes (CPP) around the time of the response (Feuerriegel et al., 2022; Kelly & O'Connell, 2013; O'Connell et al., 2012). After EEG processing, our final response-locked EEG dataset consisted of an average of 280 out of 288 epochs retained per participant (range: 246–288 for time-domain and 233–288 for frequency-domain analyses).

2.5 Behavioural Analyses

We first assessed whether RTs varied between voluntary and forced decisions. Longer voluntary RTs would indicate that participants engaged with the task appropriately, as choosing between multiple options should take longer than single-option, forced-choice decisions. Although the subsequent CoM responses are not the primary focus of this study, we also report CoM RTs for additional validation, as switching choices would be expected to produce slower RTs.

For the initial decision phase, we analysed responses made within 200–3700 ms of the stimulus. We modelled RTs using Gamma family (log-link) generalised linear mixed-effects models (GLMMs) using the *lme4* package in R (Bates et al., 2015), with decision type (voluntary vs forced) as a fixed effect and random intercepts for participants to account for individual variability:

$$RT \sim 1 + DecisionType + (1 \mid pID) \tag{1}$$

We chose a Gamma-log model because RT distributions were right-skewed and strictly positive, violating Gaussian model assumptions but making them well-suited to Gamma family models (Lo & Andrews, 2015).

We used likelihood ratio tests to compare full vs. reduced models, and to assess conditional effects. Estimated marginal means were computed using the *emmeans* v2.0 package in R (Lenth & Piaskowski, 2025). Pairwise contrasts were reported as ratios of means with 95% confidence intervals.

For CoM decisions, we fit separate GLMMs to compare switch vs. stay RTs for voluntary and forced decisions. As with the initial decision analysis, we treat CoM decision type (switch vs. stay) as a fixed effect with participant intercepts as random effects.

2.6 EEG Analyses

We analysed response-locked data to assess accumulation-to-bound dynamics corresponding to evidence accumulation and motor preparation before the response. We measured the CPP at electrode Pz, as done in Feuerriegel et al. (2021), Steinemann et al. (2018), and Tagliabue et al. (2019) to index evidence accumulation, and both MB amplitudes and the LHRP at electrode C3 (left motor cortex, contralateral to the right hand) to index motor preparation.

To derive MB time-frequency amplitudes estimates, we applied complex Morlet wavelet convolution to the CSD-transformed data (1–30 Hz in 1 Hz steps, with wavelet cycles increasing linearly from 3 to 10 across the frequency range). Time-frequency amplitude estimates were converted to decibels (dB) relative to the median amplitude across all trials for each participant, using a baseline from -500 ms to -300 ms relative to

stimulus onset to minimise contributions from the fixation stimulus presentation. We then created response-locked epochs from the dB-transformed, baseline-corrected data. MB amplitudes were averaged across the 8–30 Hz range (as done in e.g., Feuerriegel et al., 2021; Steinemann et al., 2018; Twomey et al., 2016) at electrode C3.

2.6.1 Characterising Evidence Accumulation and Motor Preparation Signatures of Voluntary Decisions

We tested whether the CPP, MB, and LHRP exhibited hallmark dynamics of evidence accumulation during the initial decision phase: (i) pre-response slopes scaling inversely with RTs (consistent with temporal integration of decision evidence), and (ii) pre-response amplitudes converging across RTs (consistent with a fixed decision threshold).

We defined pre-response amplitudes as the mean signal from –130 ms to –70 ms relative to the response, capturing immediate pre-response activity while accounting for an estimated 70 ms motor delay between decision and action execution (Kelly et al., 2021), and avoiding earlier amplitude dynamics associated with changes in build-up rates (Feuerriegel et al., 2022). We estimated slopes by fitting a linear regression model to each single-trial response-locked epoch from –300 to –70 ms, in line with prior work (Feuerriegel et al., 2021; Steinemann et al., 2018).

To relate RTs to amplitude and slope measures, we fit linear mixed-effects models (LMMs) with these measures as the outcome and RT as the predictor. Voluntary RTs are slower, more skewed and more variable than for forced decisions, with notable scale differences between participants. To account for this, we log-transformed RTs to reduce skew and z-scored them within participants to focus coefficients on within-participant variance (Baayen et al., 2008; Lo & Andrews, 2015). Models included participant intercepts and, where possible, random slopes for RT, allowing the strength of the RT–ERP measure relationship to vary across participants (Barr, 2013):

$$ERP_{measure} \sim 1 + RT + (1 + RT \mid pID) \tag{2}$$

If the maximal model – the random-slopes model – produced a singular fit or failed to converge (e.g., due to limited trials, near-zero variance components, or collinearity), we removed the random slope and retained a random intercept (Bates et al., 2015). We compared models with and without the fixed effect of RT (within matched random effects structures) using maximum likelihood and likelihood-ratio tests, and we report regression coefficients (β) from the best-fitting converged model. Results for likelihood ratio tests and regression coefficients are reported in the Supplementary Materials Tables 1–4.

To test for RT-related effects beyond a priori defined time windows, we performed mass-univariate analyses across the response-locked epochs (as in Kelly and O'Connell, 2013) with cluster-based permutation corrections for multiple comparisons (Maris & Oostenveld, 2007) implemented in the Decision Decoding Toolbox (Bode et al., 2019). For each trial and time point, we computed CPP, MB, and LHRP amplitudes and estimated linear slopes in 300 ms windows (as per minimum window width recommendations from Cohen, 2019 for effective temporal resolution of Morlet MB estimates) centred on each time point. For slope estimates, we excluded the first and last 150 ms where a full 300 ms window could not be created.

We then fit regression models to predict slope and amplitude values at each time point using RT as a predictor. One-sample t-tests were performed to compare the distributions of the resulting β values against zero. We formed clusters from neighbouring time points with $p < .01$ (uncorrected), and summed t-values to compute cluster masses. We permuted condition labels for a subset of participants (10,000 permutation iterations) to build an empirical null distribution using the largest observed cluster mass per permutation. Clusters exceeding the 97.5^{th} percentile were deemed significant ($\alpha = .05$,

two-tailed), controlling for the weak family-wise error rate (Groppe et al., 2011; Maris & Oostenveld, 2007). Surviving clusters denoted time windows where amplitudes and slopes covaried with RT.

Using the BayesFactor toolbox v3.0 (Krekelberg, 2024), we also derived Bayes Factors (BF_{10}) from Bayesian one-sample *t*-tests with default Cauchy priors (r = 0.707; Rouder et al., 2017). The BF_{10} is a ratio that reflects the likelihood of observing the data under the alternative hypothesis as compared to the null. We interpret BF_{10} as follows: $BF_{10} > 3$ positive, $BF_{10} > 20$ strong, and $BF_{10} > 150$ very strong preferential evidence for the alternative hypothesis (Kass & Raftery, 1995).

Finally, we compared group-averaged waveform amplitudes of the three signals between voluntary and forced decisions using mass-univariate paired-samples *t*-tests, using the same cluster-based permutation correction as above. This allowed us to determine whether voluntary and forced decisions produced any systematic differences in the morphologies of each signal.

3 Results

To validate our task manipulation, we tested whether RTs differed across decision types: initial forced vs voluntary decisions, and subsequent switch vs stay decisions. Individual and group mean RTs for each decision type are plotted in Figure 2A. Initial forced decisions (M = 0.86 s) were faster than voluntary decisions (M = 1.04 s; $\chi^2(1)$ = 472.59, $p < .001$, ratio = 0.83, 95% CI [0.81, 0.84]). Subsequent switch decisions were slower than stay decisions, suggesting a switch cost associated with decision updating. This was the case for voluntary switch (M = 1.29 s) and stay decisions (M = 1.23 s; $\chi^2(1)$ = 9.11, p = .003, ratio = 0.95, 95% CI [0.92, 0.98]) and forced switch (M = 1.12 s) and stay decisions (M = 1.03 s; $\chi^2(1)$ = 28.26, p < .001, ratio = 0.92, 95% CI [0.90, 0.95]). Crucially, RTs for all forced decisions were slower than typical simple reaction time ($\sim$ 235

ms; Woods et al., 2015) and discrimination time ($\sim$ 150 ms; Hopf et al., 2002) latencies, indicating that forced decisions entailed colour identification and decision commitment rather than reflexive detection. Together, these RT patterns suggest participants engaged with both the initial and subsequent decision task as instructed.

3.1 Characteristics of the CPP, MB, and LHRP Signals

Pre-response amplitude scalp maps for both voluntary and forced decisions in Figure 2B showed a prominent centro-parietal positivity focused at Pz, which was spatially distinct from motor-related activity over the left motor cortex at C3. MB scalp topographies in Figure 2C also showed focal power reduction (i.e., more negative dB values) over left-central sites near C3 for both decision types. Importantly, in this CSD-transformed data, a typical RP-like negativity focused over Cz was not apparent in this single-effector response design. Across voluntary and forced decisions, all three neural signals exhibited the expected ramping morphology leading up to the motor response: a positive-going CPP build-up (Figure 2D) and negative-going build-ups for MB amplitudes (Figure 2E) and the LHRP (Figure 2F).

To test for any clear differences across decision types, we compared voluntary vs forced waveform amplitudes using mass-univariate analyses. Across all three signals, we did not find reliable waveform differences within time windows immediately prior to the response. For MB amplitudes, an early significant cluster emerged between -740 and -460 ms (cluster $p < .001$). This window overlapped with a period where CPP BFs indicated moderate support for amplitude differences but this did not survive multiple testing corrections. These early effects likely reflect ramp-onset timing differences driven by RT differences between voluntary and forced decisions, as shown in Figure 2A and lay outside our pre-response window of interest. We also found moderate BFs in support of MB amplitude differences around response onset (-100 to -10 ms), but these effects also did not

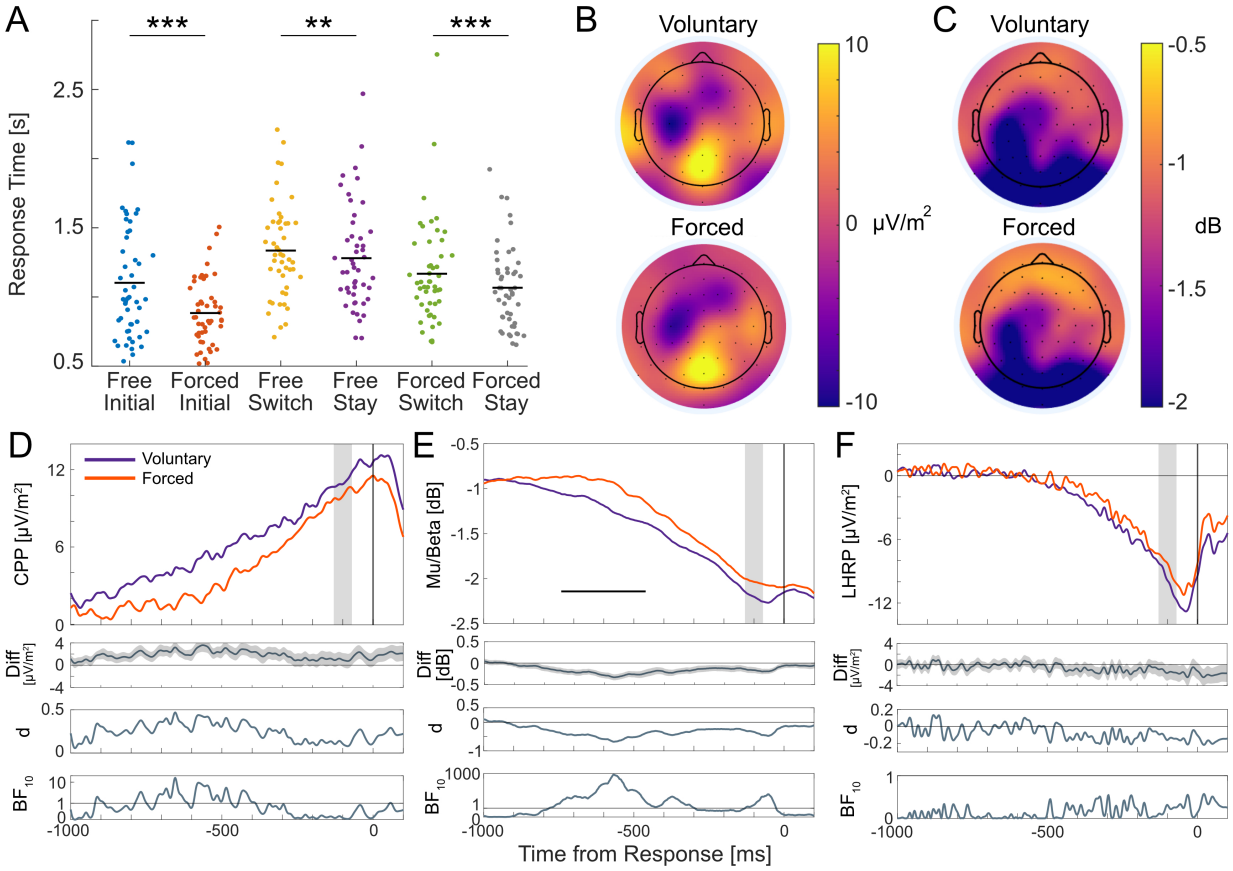


Figure 2

*Mean response times across decisions and EEG signal dynamics. A) Participant (dots) and group (horizontal black lines) mean response times by decision type. Lines with asterisks denote statistically significant response time differences (** denotes $p < .01$, *** denotes $p < .001$). B - C) Response-locked scalp topographies of voluntary (top) and forced (bottom) decisions averaged across the -130 to -70 ms window preceding response onset. Panel B shows time-domain ERP amplitudes; Panel C shows Mu/Beta (8–30 Hz) spectral amplitudes. D) Group-averaged CPP, E) MB amplitudes, and F) LHRP for voluntary (purple) and forced (orange) waveforms with accompanying difference waves (and shaded standard error regions), effect size estimates (Cohen, 1988), and BFs in favour of the alternative hypothesis (i.e., amplitude differences) at each time point. Grey areas represent the -130 to -70 ms window used to derive the scalp maps. The horizontal black line under MB waveforms denotes statistically significant differences after correcting for multiple comparisons. Horizontal lines in BF plots denote $BF_{10} = 1$, indicating a lack of preferential support for either the alternative or null hypothesis. Overall, all three signals exhibited the expected ramping morphology, with no reliable amplitude differences between voluntary and forced decisions immediately before the response.*

survive correction for multiple comparisons. The LHRP did not show statistically significant amplitude differences at any time point. Given these canonical profiles, we

proceeded to test accumulation-to-bound dynamics across the three signals.

3.2 The CPP exhibits accumulation-to-bound dynamics in voluntary decisions

We examined whether the CPP, MB amplitudes, and the LHRP exhibited accumulation-to-bound characteristics: (i) build-up slopes that scaled inversely with RTs, and (ii) pre-response amplitudes that converged to similar levels irrespective of RTs. These dynamics were assessed in two ways: first, using fixed pre-response windows for slopes (-300 to -70 ms) and amplitudes (-130 to -70 ms) as defined in prior perceptual decision-making work (e.g., Feuerriegel et al., 2021; O'Connell et al., 2012; Steinemann et al., 2018); second, using mass-univariate analyses across the full response-locked epoch to identify any additional RT-related effects outside the pre-defined windows. Here, we present data relative to the initial decision in each trial. Data relating to CoM decisions are beyond the scope of our paper.

To visualise these results, Figures 3A and B show group-averaged stimulus-locked (left) and response-locked (right) waveforms for voluntary and forced decisions, respectively, split by RT bins (fast, medium, and slow). Trials were split into three equal-sized RT bins for each participant and averaged across participants. Please note that, as RT distributions for both voluntary and forced decisions were positively skewed and could extend up to 3.7s due to the self-paced nature of the task, the slowest RT bins spanned a wider RT range and may show greater temporal variability than the fastest RT bin. This may accentuate apparent visual differences across the RT bins, such as around the pre-response amplitude window in Figures 3A and B (right panel), and Figure 5B, although shading around the waveforms denoting standard errors overlaps slightly across bins within this window. Thus, visual differences across RT bins should be interpreted with some caution and evaluated against the inferential statistics reported below, which were performed on single-trial EEG data.

Response-locked waveforms in Figures 3A and B (right panel) show the expected gradual build-up across RT bins – with steeper build-ups for the fastest RT bin – and amplitudes converging to similar pre-response levels. We also include stimulus-locked waveforms at Pz in Figures 3A and B (left panel) for completeness and descriptive context, but we do not analyse them further here. In our task, the visual display changed immediately after a keypress response. Therefore, stimulus-locked waveforms capture a combination of pre- and post-decisional ERPs, including visual evoked responses, whose timing varies with RT. Accordingly, stimulus-locked plots should be interpreted descriptively and not as a direct readout of decision dynamics.

Within the pre-defined analysis windows, faster RTs were associated with steeper CPP slopes for both voluntary ($\beta = -0.02$, $SE = 0.004$, $t(48.29) = -4.03$, $p < .001$) and forced ($\beta = -0.02$, $SE = 0.001$, $t(41.81) = -3.34$, $p = .002$) decisions. Negative coefficients indicate steeper build-up rates for faster RTs and shallower slopes for slower RTs. Mass-univariate analyses also revealed RT effects that overlapped with these pre-defined windows. Voluntary RTs covaried with CPP slopes from approximately -160 to +115 ms (cluster p < .001), and forced RTs covaried from –230 to -30 ms (cluster p = .002) and at an earlier window spanning –900 to -300 ms (cluster p < .001). BFs indicated very strong support for RT-slope associations during these windows. Together, these findings indicate that CPP slopes scale systematically with RTs in both voluntary and forced decisions, consistent with the CPP indexing an evolving decision variable.

For CPP amplitudes, including RTs as a fixed effect did not improve model fit for either voluntary decisions ($\chi^2(1) = 1.30$, $p = .254$) or forced decisions ($\chi^2(1) = 2.56$, $p = .109$). Mass-univariate analyses identified RT-amplitude associations for voluntary decisions, spanning -1050 to –555 ms (cluster p = .013) and -485 to -340 ms relative to the response (cluster p < .05); and for forced decisions from –635 to –115 ms (cluster p = .004). These effects did not extend into the time period just preceding response onset, suggesting

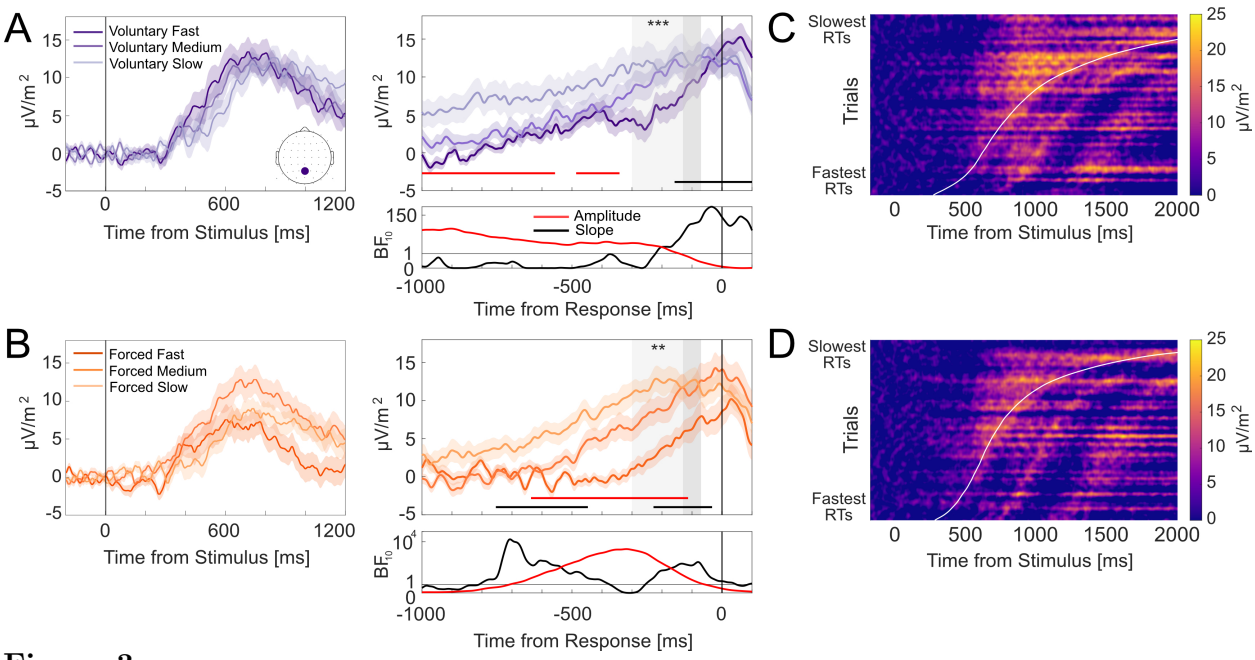


Figure 3

*CPP evidence accumulation dynamics at electrode Pz. A-B) Group-averaged stimulus-locked (left) and response-locked (right) ERP waveforms for voluntary (A) and forced (B) decisions split by fast, medium, and slow RT bins, with shading denoting standard errors. Dark grey areas denote the -130 to -70 ms pre-response amplitude window. Light grey areas denote the -300 to -70 ms time slope window. Asterisks denote statistically significant RT associations with ERP amplitudes and slope differences within their respective windows (** denotes $p < .01$; *** denotes $p < .001$). Red and black horizontal lines denote windows where RTs were associated with ERP amplitudes and slope differences, respectively. Corresponding BF_{10} estimates for each association are plotted below. Horizontal lines in BF plots denote $BF_{10} = 1$, indicating a lack of preferential support for either the alternative or null hypothesis. Bayes factors for forced decisions are plotted on logarithmic scales due to the wide ranges of values across the time-course. C-D) Single-trial stimulus-locked CPP amplitudes for voluntary (C) and forced (D) decisions pooled across participants. Trials are sorted by RT (fastest below and slowest on top). Data were smoothed over 300-trial bins using a sliding Gaussian-weighted moving average filter. RTs are denoted by the curved white line.*

that CPP amplitudes converged to similar levels around the time of the motor response.

Taken together, CPP slopes scaled reliably with RTs, while pre-response amplitudes converged to similar levels irrespective of RTs. To further visualise these accumulation-to-bound dynamics, we plotted heatmaps of single-trial stimulus-locked CPP amplitudes at electrode Pz, sorted by ascending RTs and pooled across participants for

voluntary (Figure 3C) and forced decisions (Figure 3D; as done in O'Connell et al., 2012 and Feuerriegel et al., 2022). These heatmaps provide an intuitive representation of tracing how amplitude build-up unfolds as a function of RTs, reflecting temporal integration of evidence to a decision bound. For both voluntary and forced decisions, the heatmaps show a gradual positive-going build-up in CPP amplitudes (i.e., increasing yellow saturation) at similar onsets across trials that culminate at comparable peak levels around the time of the response (denoted by the curved black line).

3.3 MB amplitudes exhibit accumulation-to-bound dynamics in voluntary decisions

Response-locked MB amplitude trajectories for voluntary and forced decisions, split by fast, medium, and slow RT bins, are displayed in Figures 4A and B, respectively. Both figures show MB amplitudes generally converged to a common value prior to the response and the rate of amplitude decrease was more gradual for slower RTs. For MB slopes in the pre-defined window, including RTs as a predictor did not improve model fit for voluntary ($\chi^2(1) = 2.16$, $p = .141$) or forced ($\chi^2(1) = 0.01$, $p = .909$) decisions. However, outside of this window, mass-univariate analyses detected significant RT effects spanning -450 to -345 ms for voluntary decisions (cluster p = .039), and between -915 and -730 ms (cluster p = .004) and -365 to -50 ms (cluster p < .001) for forced decisions. This indicates that for both voluntary and forced decisions, faster RTs were accompanied by steeper (i.e., more negative) MB slopes, whereas longer RTs showed shallower (i.e., less negative) slopes. These results suggest that the pre-defined time window (based on existing work measuring the CPP) was not well-suited to detecting slope differences in MB amplitudes, which exhibit substantially more temporal smearing than ERPs due to the wavelet convolution procedure. The timing of the statistically significant clusters is entirely consistent with RT-dependent slope differences in MB trajectories.

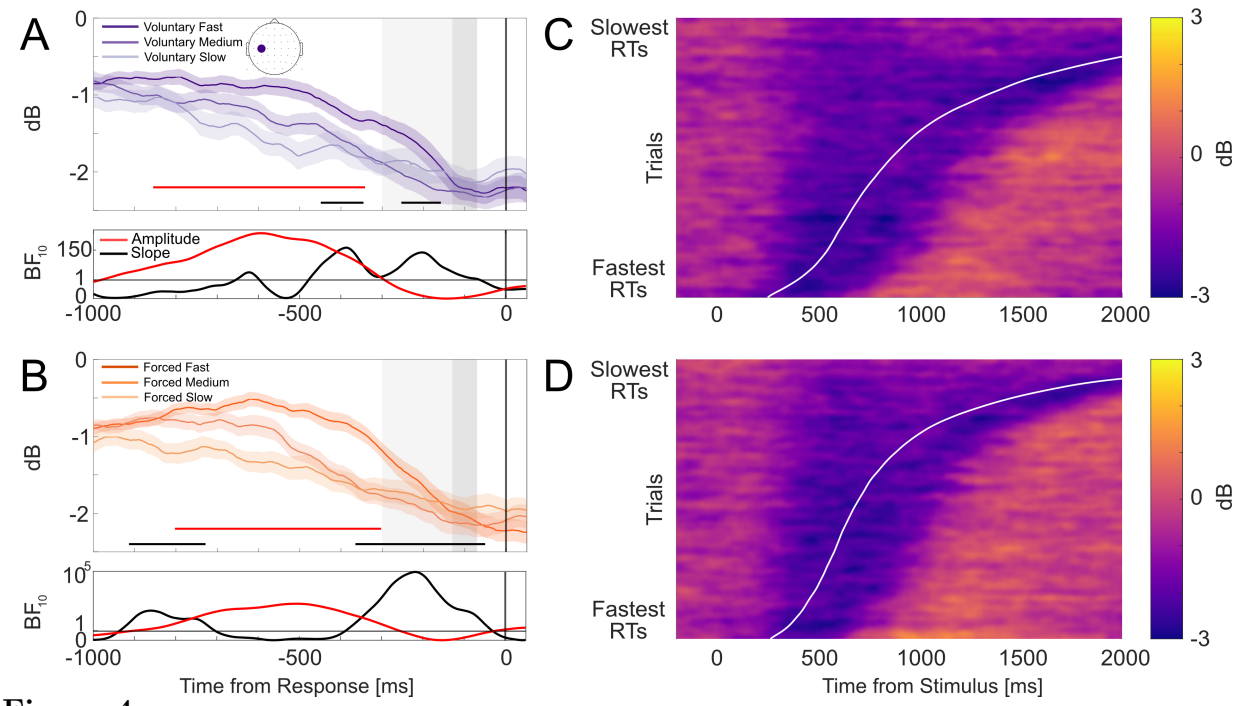


Figure 4

Mu/Beta (8–30 Hz) dynamics at C3 (scalp map in top-left panel). A-B) Group-averaged response-locked waveforms for voluntary (A) and forced (B) decisions split into fast, medium, and slow RT bins, with shading denoting standard errors. Dark grey areas denote the -130 to -70 ms pre-response amplitude window. Light grey areas indicate the -300 to -70 ms slope window. Red and black horizontal lines denote windows where RTs were associated with amplitude and slope differences, respectively. Corresponding BF_{10} estimates for each association are plotted below. Horizontal lines in BF plots denote $BF_{10} = 1$, indicating a lack of preferential support for either the alternative or null hypothesis. Note that for (B), BFs are plotted on logarithmic scales due to the wide ranges of values across the time-course. C-D) Heat maps depicting MB amplitudes for voluntary (C) and forced (D) decisions relative to stimulus onset. Trials are sorted by RT (fastest below and slowest on top). Data were smoothed over 300-trial bins using a sliding Gaussian-weighted moving average filter. RTs are denoted by the curved white line.

For MB amplitudes, adding RTs as a fixed effect did not improve model fit for voluntary ($\chi^2(1) = 0.19$, $p = .664$) or forced ($\chi^2(1) = 0.56$, $p = .452$) decisions. Similarly, mass univariate analyses did not reveal support for RT-amplitude effects, particularly in the 400 ms interval before the response.

Heatmaps of single-trial MB amplitudes for voluntary (Figure 4C) and forced (Figure 4D) decisions both show accumulation-to-bound dynamics closely resembling those

of the CPP: gradual decrease in amplitudes, indexed by increasingly purple saturation, beginning at comparable onsets across trials and peaking at similar amplitude levels around the response.

3.4 The LHRP exhibits threshold-crossing dynamics but may not trace evidence accumulation

Figures 5A and B show response-locked group-averaged LHRP waveforms for voluntary and forced decisions, respectively, split into fast, medium, and slow RT bins. Both figures show LHRP waveforms that gradually became more negative-going and reached their peak just prior to the response. Although forced decision waveforms in Figure 5B display visual separation across RT bins (like in the CPP), these separations are likely accentuated by the highly skewed RT distributions. These waveforms should be interpreted with the same caution noted for the CPP.

Within the pre-defined analysis window, RTs significantly predicted LHRP slopes for voluntary decisions ($\beta = 0.01$, $SE = 0.005$, $t(49.9) = 2.42$, $p = .019$), indicating that slower RTs were associated with shallower LHRP slopes. However, mass-univariate analyses did not reveal any reliable RT effects at any time point. As cluster-based permutation tests are more sensitive to detecting effects that span across several consecutive time points, the absence of significant clusters found suggests that RT effects found during the pre-defined window may instead reflect local variability or effects limited to that isolated time window. Furthermore, BFs across time points within the pre-defined window were less than 1, indicating a lack of clear evidence for RT-slope effects. For forced decisions, including RTs as a fixed effect did not improve model fit ($\chi^2(1) = 1.06$, $p = .303$). Mass univariate analyses and BFs also did not provide substantial evidence in support of reliable RT-slope effects at any time point.

For the analysis of LHRP amplitudes, including RTs as a fixed effect did not

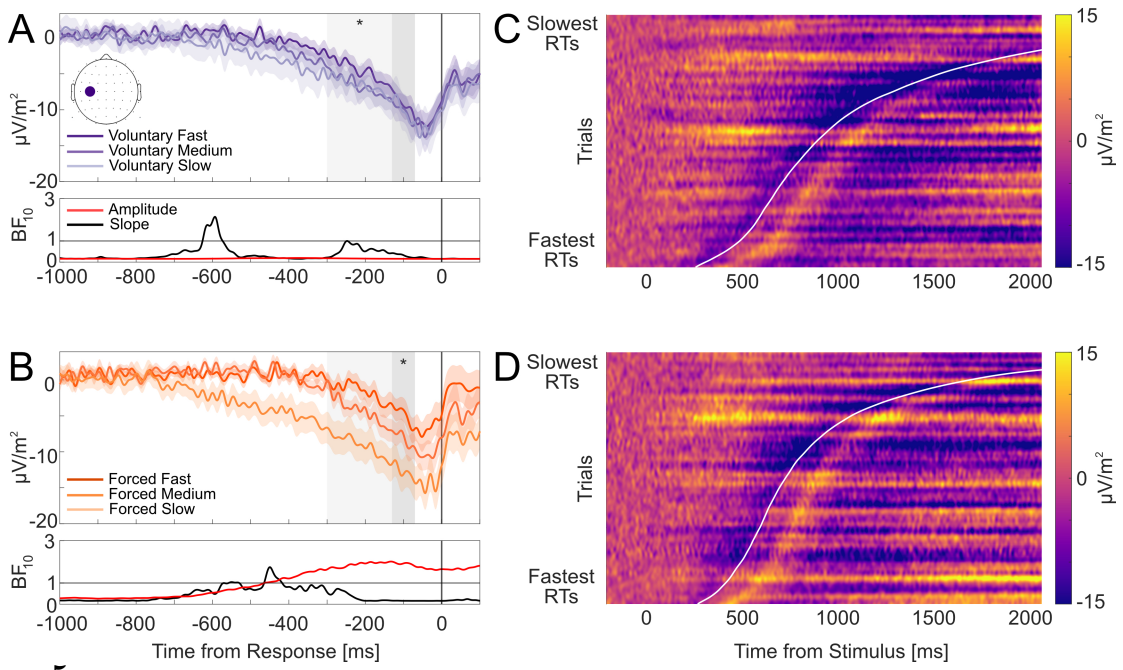


Figure 5

Left hemisphere readiness potentials at electrode C3 (scalp map in top-left panel). A-B) Group-averaged response-locked waveforms for voluntary (A) and forced (B) decisions split into fast, medium, and slow RT bins, with shading denoting standard errors. Dark grey areas denote the -130 to -70 time pre-response amplitude window. Light grey areas indicate the -300 to -70 ms slope window. Asterisks denote statistically significant RT associations between ERP amplitudes and slope differences within their respective windows (denotes p < .05). Corresponding BF_{10} estimates in red and black denote preferential support for RT associations with amplitude and slope differences, respectively, at each time point are plotted below. Horizontal lines in BF plots denote $BF_{10} = 1$, indicating a lack of preferential support for either the alternative or null hypothesis. However, we did not observe any significant clusters in which RT modulations on the amplitudes and slopes were sustained over time. C-D) Heat maps depicting LHRP amplitudes for voluntary (C) and forced (D) decisions relative to stimulus onset. Trials are sorted by RT (fastest below and slowest on top). Data were smoothed over 300-trial bins using a sliding Gaussian-weighted moving average filter. RTs are denoted by the curved white line.*

improve model fit for voluntary decisions ($\chi^2(1) = 0.32$, $p = .571$). In contrast, for forced decisions, RT was a significant predictor of amplitudes ($\beta = -3.53$, $SE = 1.34$, $t(45.37) = -2.62$, $p = .012$), with slower RTs associated with more negative-going amplitudes. However, mass-univariate analyses and BFs did not provide preferential support for RT-amplitude associations at any time point for voluntary decisions nor substantial support for forced decisions ($BF_{10} < 3$), which may similarly indicate that

effects during the pre-defined time window were not sustained for long enough beyond the localised time window or sufficiently strong to survive corrections for multiple comparisons.

The heatmaps for voluntary (Figure 5C) and forced (Figure 5D) decisions both show negative-going build-up that peaks at similar amplitude levels around response onset regardless of RT. However, the LHRP build-up begins at a relatively fixed time before the response in contrast to earlier build-ups observed in the CPP and MB, reflecting dynamics that align with a late-stage, effector-specific motor preparation rather than extended evidence integration.

Taken together, the analyses do not yield strong support for RT modulation of LHRP slopes or pre-response amplitudes. However, the visual patterns in Figure 5B and localised effects in the pre-defined slope and amplitude windows leave open some possibility for subtle RT-related influences that may have been too weak to detect in our data.

4 Discussion

We assessed whether neural signatures of evidence accumulation and motor preparation in perceptual decision tasks also characterise endogenously guided, voluntary decisions, which are primarily driven by abstract goals and preferences. We also examined forced decisions whereby a single colour specified the response, similar to perceptual decision-making designs (such as colour identification tasks). Importantly, our experimental design allowed us to investigate the formation of voluntary decisions independently from action selection or effector choice.

After accounting for stimulus-locked contributions to response-aligned neural activity, waveform amplitudes and pre-response scalp topographies of CPP, MB and LHRP signals were highly similar between voluntary and forced decisions. Furthermore, both CPP and MB signals displayed accumulation-to-bound properties: (i) build-up rates

(slopes) were steeper for faster RTs and (ii) pre-response amplitudes converged to similar levels around the time of the motor response. The LHRP findings were more mixed: we observed an RT–slope association for voluntary decisions and an RT-amplitude association for forced decisions within their respective pre-response measurement windows. However, these effects were not clearly apparent in our mass-univariate analyses and did not survive correction for multiple comparisons. Therefore, our results suggest that the LHRP ramps towards a stereotyped amplitude level before a motor response, but may not trace evidence accumulation in a similar manner to CPP and MB signals. Taken together, these findings point to shared evidence accumulation and motor preparatory mechanisms that underlie internally and externally guided choices.

4.1 The CPP traces evidence accumulation during both perceptual and internally-driven voluntary decisions

Recent work has questioned whether the CPP exclusively indexes evidence accumulation in perceptual decision-making, given the potential influence of measurement confounds in prior value-based judgement tasks (Frömer et al., 2024). When accounting for these confounds, we observed canonical accumulation-to-bound CPP dynamics in voluntary decisions that embodied a value-based decision – deliberation between two colours based on internally generated preferences and goals. O'Connell et al. (2025) cautioned that true CPP-like dynamics in Frömer et al. (2024) may have been masked by methodological and analysis decisions. We mitigated these issues by applying deconvolution (RIDE; Ouyang et al., 2015) to reduce stimulus-related overlap in our response-locked data, and by using CSD transformations to enhance spatial precision of EEG signals (Feuerriegel et al., 2022; Kelly and O'Connell, 2013). The emergence of the CPP under these conditions lends support to its role as a correlate of a domain-general decision variable rather than one restricted to perceptual decisions. We also present, to our knowledge, the first evidence of CPP and MB correlates of evidence accumulation in a task commonly used to study

voluntary (i.e., intentional) decisions.

Despite qualitative differences in task demands, CPP waveform amplitudes did not differ reliably between voluntary and forced decisions, and both showed accumulation-to-bound properties. In our task, voluntary decisions required recognising that a choice existed and then deliberating between two alternatives based on internally generated preferences or goals, without the constraints of an objectively "correct" answer. By contrast, forced decisions followed a fixed rule – choose the sole presented option – but nonetheless entailed a perceptual decision to identify and commit to the colour presented. Crucially, forced decision RTs were slower than typical simple reaction or discrimination task RTs (∼150 - 235 ms; Hopf et al., 2002; Woods et al., 2015), indicating decisional processing rather than reflexive responding. A deliberate decision was required as each colour choice needed to be maintained in memory in anticipation of a possible change of mind in half of the trials. Together, these observations indicate that the CPP tracks a common decision variable irrespective of whether evidence is exogenous or endogenously generated, supporting sequential-sampling accounts of voluntary choice.

4.2 MB amplitudes trace the accumulation of decision evidence in the motor system

Like the CPP, MB amplitudes reflected an RT-dependent build-up to a fixed pre-response level for both voluntary and forced decisions, suggesting similar motor-gating policies in which actions are triggered once motor preparatory activity reaches a stereotyped threshold (Donner et al., 2009; Lange et al., 2013; O'Connell et al., 2012). We note that MB trajectories can be sensitive to urgency effects under speeded response deadlines. MB amplitudes can progress toward the response-triggering threshold at a faster rate than expected based on evidence accumulation rates, and can even begin ramping prior to the onset of a decision-relevant stimulus (Kelly et al., 2021; Steinemann et al.,

2018). However, our MB results are more consistent with a decision variable within the motor system where there is little speed pressure: when evidence was accumulated slowly, MB ramped more gradually and resembled the trajectories of the CPP. Variations in response onset may better be explained by variations in evidence accumulation rates rather than idiosyncratic changes in motor-execution thresholds (Gold & Shadlen, 2000; Kim & Shadlen, 1999). Furthermore, our task did not explicitly impose speed pressure manipulations – a context in which MB and CPP dynamics are often similar (e.g., Donner et al., 2009; Feuerriegel et al., 2021; Lange et al., 2013; O'Connell et al., 2012).

Our MB dynamics also align with accumulator modelling results in speeded perceptual decision tasks, which reliably show fixed boundary setting with variability in stimulus-driven (e.g., ease of discrimination) or goal-driven (e.g., reward rate) evidence accumulation rates (Brown & Heathcote, 2008; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). These dynamics possibly reflect the lack of urgency manipulations and do not support the idea that urgency is an essential feature of voluntary decisions as previously proposed by Brass et al. (2019). Future work could introduce urgency manipulations (e.g., deadlines) to test whether neural signatures of urgency – reduced CPP and stable MB pre-response amplitudes for slower decisions – are observed (Feuerriegel et al., 2021; Kelly et al., 2021; Steinemann et al., 2018).

Having identified neural correlates of evidence accumulation (i.e., CPP and MB signals) in voluntary and forced decisions – with broadly similar waveform amplitudes across both decisions – our findings further support extending sequential sampling models from perceptual decision-making to voluntary, multi-alternative decisions. This motivates future joint neural-behavioural modelling (e.g., hierarchical DDM and race models) of voluntary decisions (as done in Kelly and O'Connell, 2013 for perceptual decisions and Sun et al., 2025 for recognition memory decisions). For example, CPP slopes can serve as a proxy for trial-wise drift rates to link variability in RTs to endogenous evidence, such as

individual choice preferences, and CPP pre-response amplitudes could map onto decision-bound criteria. MB signals can be modelled similarly in voluntary decisions made under speed pressure to assess whether MB signals reflect urgency-related boundary shifts – indexed by flexible decision bound criterion and RT-amplitude associations – and whether the CPP consistently tracks evidence accumulation trajectories, as in perceptual tasks (Kelly et al., 2021; Steinemann et al., 2018). Such modelling efforts will advance models of voluntary decisions beyond "when to act" (e.g., Schurger, 2018; Schurger et al., 2012) to "which option" decision dynamics, and further clarify the functional commonalities that underlie evidence accumulation in voluntary and forced decisions.

4.3 The LHRP indexes late motor preparation but may not reflect evidence accumulation dynamics

Relative to the CPP and MB amplitudes, the LHRP showed less consistent evidence for accumulation-to-bound trajectories for both voluntary and forced decisions. We observed subtle RT-slope associations in the pre-response window for voluntary decisions. Conversely, in forced decisions, slower RTs were associated with more negative LHRP amplitudes in the –130 to –70 ms window, but this effect was not observed for voluntary decisions. Although this RT-amplitude association could be interpreted as increasing boundaries with elapsed time, the directionality of this pattern in our results runs counter to more commonly-proposed collapsing-bound predictions (e.g., COINTOB model; Brass et al., 2019), where slower responses terminate at lower thresholds (O'Connell & Kelly, 2021; Steinemann et al., 2018). Moreover, these effects were not statistically significant in the mass-univariate analyses after correcting for multiple comparisons.

Although accumulation-like trajectories have previously been reported for the LRP using analyses similar to our study (Kelly & O'Connell, 2013), and for the RP using computational modelling (Lui et al., 2021; Maoz et al., 2019), these measures differ from

our LHRP in critical ways that limit their direct comparison. The classical RP is measured at Cz, indexing broader central motor preparatory activity, thought to originate from pre-SMA and the SMA. However, without CSD transformations, activity at Cz can capture overlapping signals from neighbouring sites, such as motor-independent CPP signals from Pz. The LRP is computed as the contralateral–ipsilateral difference between C3 and C4, indexing bilateral effector competition. LRP ramping trajectories arise from participants choosing between hands, yielding a growing lateralised difference in motor preparation. Meanwhile, in our single-effector design, we mapped every decision to a right-handed keypress, thereby eliminating effector competition and motivating a focus on contralateral activity at C3. This choice is further supported by the scalp maps in Figure 2B, which also did not show an RP at Cz in our CSD-transformed data. Therefore, it is plausible that previously reported evidence accumulation dynamics may in part arise from activity related to action selection.

While we cannot conclusively rule out accumulation-to-bound dynamics here, our results suggest that when action selection is controlled for, the LHRP at C3 may be better interpreted as a late-stage effector-specific gate that releases an action once a relatively fixed motor threshold is crossed. It appears less clearly as an evolving decision variable for the motor system.

4.4 Limitations and future directions

Although our single-effector design strengthened our inferences about decision formation independent of hand selection, it also limits direct comparisons with existing RP and LRP findings that show accumulation-like dynamics in dual-handed task designs with effector competition (e.g., Kelly and O'Connell, 2013; Maoz et al., 2019), and in single-effector designs using, for example, shoulder rotation responses in different directions (Lui et al., 2021). Future studies could reintroduce effector competition using dual-handed

response designs to test whether the RP or LRP signals (guided by CSD-transformed scalp topographies) show accumulation-to-bound dynamics for voluntary decisions using the same analysis framework. Such analyses will clarify whether readiness potentials exhibit accumulation-to-bound properties only when effector competition is present, or whether they primarily reflect late-stage threshold-crossing properties – akin to our LHRP findings.

In our study, inferences relating to stimulus-locked ERPs were constrained by our task design. Visual display changes occurred immediately after the keypress response (e.g., fixation stimulus changes), and there were substantial differences in the degree of RT variability across the RT tertile bins. Consequently, plotted waveform amplitudes reflected different mixtures of pre-decisional build-up and post-decisional visual evoked activity as well as different degrees of temporal smearing. A design that delays post-response visual changes would permit cleaner isolation of pre-decisional stimulus-locked activity in future work (e.g., Sun et al., 2025). Nonetheless, our core inferences here were based on response-locked analyses that directly traced dynamics around decision commitment and are not compromised by these issues.

In addition to a prominent CPP and left-hemispheric motor activity, our scalp maps also showed a fronto-central negativity consistent with the contingent negative variation (CNV) ERP component (Figure 2B). The CNV has been proposed to index a movement-independent urgency signal that grows with elapsed time and can reset when actions are withheld, placing this temporally between evidence accumulation and motor execution (McCone et al., 2025). These properties suggest that the CNV may be more tightly linked to the termination of decisional evidence accumulation (e.g., through urgency parameters or collapsing bounds), rather than motor execution. At the same time, CNV trajectories closely resemble the RP in morphology and topography, although their distinct functional roles in decision-making remain unclear (see Schurger et al., 2021 for detailed discussion). We did not focus on the CNV here as our a priori questions concerned

decisional evidence accumulation (i.e., CPP) and motor preparation (i.e., MB and LHRP) trajectories. Additionally, our task did not impose explicit urgency manipulations – conditions under which CNV dynamics have been most informative. Future studies could introduce deadlines to test whether the CNV unfolds as an evolving decision variable with accumulation-to-bound trajectories like the urgency-insensitive CPP, or as a time-dependent urgency signal that lowers effective decision thresholds as time elapses (as found in McCone et al., 2025). Observing such CNV–CPP coupling under time pressure provides a decisive test of urgency-based accounts and clarifies how urgency shapes the interplay between decision termination and motor gating in voluntary decisions.

Lastly, our results extend CPP findings from perceptual tasks to voluntary, internally driven choices, including simple value-based decisions. Although this lends support to the view that the CPP acts as a domain-general correlate of evidence accumulation, previous reports of CPP-like signals in value-based (Pisauro et al., 2017) and memory-based decisions (van Ede & Nobre, 2024; van Vugt et al., 2019) have warranted careful interpretation due to methodological limitations in these studies (O'Connell et al., 2025). To validate prior CPP reports while addressing methodological concerns raised, the present analysis framework could be applied to value- and memory-based decision datasets, as well as other domains (e.g., food preference and moral decision-making) to more decisively adjudicate whether the CPP consistently tracks evidence accumulation beyond perceptual and voluntary domains.

4.5 Conclusion

This study extends our current understanding of how voluntary decisions evolve and are converted into action. Despite prior reports of anatomical and qualitative differences between voluntary and forced actions, we provide neural evidence that voluntary decisions, like forced decisions, unfold as a noisy evidence accumulation process that terminates at a

stereotypical boundary. Crucially, the presence of canonical CPP dynamics in voluntary and simple value-based decisions suggests that the CPP is not exclusive to perceptual decision making. The LHRP did not show clear evidence of an evolving decision signal under matched motor demands, and may instead reflect late effector-specific preparation. Together, these findings point to common decision commitment mechanisms and motor-gating mechanisms between internally and externally driven decisions.

Data and Code Availability Statement

All code and data used for the experiment and analyses described in this manuscript are available at http://osf.io/t7z4h.

CRediT authorship contribution statement

Lauren C. Fong: Writing – review & editing, Writing – original draft, Visualisation, Methodology, Formal analysis, Data curation, Conceptualisation. **Paul M. Garrett**: Writing – review, & editing, Formal analysis. **Philip L. Smith**: Writing – review & editing, Supervision. **Robert Hester**: Writing – review & editing, Supervision. **Stefan Bode**: Writing – review & editing, Supervision, Funding acquisition, Conceptualisation. **Daniel Feuerriegel**: Writing – review & editing, Visualisation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualisation.

Funding

This project was supported by an Australian Research Council Discovery Early Career Researcher Award to D.F. (ARC DE220101508) and the Melbourne Research Scholarship awarded to L.C.F from the University of Melbourne. Funding sources had no role in study design, data collection, analysis or interpretation of results.

Declaration of Competing Interests

None

Ethical statement

The experiment was approved by the Human Research Ethics Committee (ID 27017) of the University of Melbourne.

Acknowledgements

We thank our laboratory members, Jie Sun in particular, and Redmond O'Connell, Simon Kelly, Elaine Corbett, Elisabeth Parés-Pujolràs, and Silvia Seghezzi for helpful discussions relating to this paper.

References

Adam, K. C. S., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive psychology*, *97*, 79. https://doi.org/10.1016/J.COGPSYCH.2017.07.001

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. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00328

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bode, S., Bogler, C., & Haynes, J.-D. (2013). Similar neural mechanisms for perceptual guesses and free decisions. *NeuroImage*, *65*, 456–465. https://doi.org/10.1016/j.neuroimage.2012.09.064

Bode, S., Feuerriegel, D., Bennett, D., & Alday, P. M. (2019). The Decision Decoding ToolBOX (DDTBOX) – A Multivariate Pattern Analysis Toolbox for Event-Related Potentials. *Neuroinformatics*, *17*(1), 27–42. https://doi.org/10.1007/s12021-018-9375-z

Bode, S., Murawski, C., Soon, C. S., Bode, P., Stahl, J., & Smith, P. L. (2014). Demystifying "free will": The role of contextual information and evidence accumulation for predictive brain activity. *Neuroscience and Biobehavioral Reviews*, *47*, 636–645. https://doi.org/10.1016/j.neubiorev.2014.10.017

Bogler, C., Grujičić, B., & Haynes, J.-D. (2023). Clarifying the nature of stochastic fluctuations and accumulation processes in spontaneous movements [Publisher: Frontiers]. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1271180

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brass, M., Furstenberg, A., & Mele, A. R. (2019). Why neuroscience does not disprove free will. *Neuroscience and Biobehavioral Reviews*, *102*, 251–263. https://doi.org/10.1016/j.neubiorev.2019.04.024

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. https://doi.org/10.1016/j.cogpsych.2007.12.002

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

Cohen, M. X. (2019). A better way to define and describe Morlet wavelets for time-frequency analysis. *NeuroImage*, *199*, 81–86. https://doi.org/10.1016/j.neuroimage.2019.05.048

Cui, R. Q., Huter, D., Lang, W., & Deecke, L. (1999). Neuroimage of Voluntary Movement: Topography of the Bereitschaftspotential, a 64-Channel DC Current Source Density Study. *NeuroImage*, *9*(1), 124–134. https://doi.org/10.1006/nimg.1998.0388

Deecke, L., Scheid, P., & Kornhuber, H. H. (1969). Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Experimental Brain Research*, *7*(2), 158–168. https://doi.org/10.1007/BF00235441

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dirnberger, G., Lang, W., & Lindinger, G. (2008). A new method to determine temporal variability in the period of pre-movement electroencephalographic activity. *International Journal of Psychophysiology*, *70*(3), 165–170. https://doi.org/10.1016/j.ijpsycho.2008.08.006

Donner, T. H., Siegel, M., Fries, P., & Engel, A. K. (2009). Buildup of Choice-Predictive Activity in Human Motor Cortex during Perceptual Decision Making. *Current Biology*, *19*(18), 1581–1585. https://doi.org/10.1016/j.cub.2009.07.066

Eimer, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 146–156. https://doi.org/10.3758/BF03209424

Feuerriegel, D., Jiwa, M., Turner, W. F., Andrejević, M., Hester, R., & Bode, S. (2021). Tracking dynamic adjustments to decision making and performance monitoring processes in conflict tasks. *NeuroImage*, *238*, 118265. https://doi.org/10.1016/j.neuroimage.2021.118265

Feuerriegel, D., Murphy, M., Konski, A., Mepani, V., Sun, J., Hester, R., & Bode, S. (2022). Electrophysiological correlates of confidence differ across correct and erroneous perceptual decisions. *NeuroImage*, *259*, 119447. https://doi.org/10.1016/j.neuroimage.2022.119447

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, *69*(3), 548–562. https://doi.org/10.1016/j.neuron.2010.11.045

Frömer, R., Nassar, M. R., Ehinger, B. V., & Shenhav, A. (2024). Common neural choice signals can emerge artefactually amid multiple distinct value signals. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-024-01971-z

Gluth, S., Rieskamp, J., & Büchel, C. (2013). Classic EEG motor potentials track the emergence of value-based decisions. *NeuroImage*, *79*, 394–403. https://doi.org/10.1016/j.neuroimage.2013.05.005

Gold, J. I., & Shadlen, M. N. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature*, *404*(6776), 390–394. https://doi.org/10.1038/35006062

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nat. Rev. Neurosci.*, *9*(12), 934–946. https://doi.org/10.1038/nrn2497

Haggard, P. (2009). The Sources of Human Volition. *New Series*, *324*(5928), 731–733. https://doi.org/10.1126/science.ll23604

Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology*, *17*(4), 323–328. https://doi.org/10.1016/j.cub.2006.11.072

Hopf, J.-M., Vogel, E., Woodman, G., Heinze, H.-J., & Luck, S. J. (2002). Localizing visual discrimination processes in time and space. *Journal of Neurophysiology*, *88*(4), 2088–2095. https://doi.org/10.1152/jn.2002.88.4.2088

Jahanshahi, M., Jenkins, I. H., Brown, R. G., Marsden, C. D., Passingham, R. E., & Brooks, D. J. (1995). Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. *Brain: A Journal of Neurology*, *118 (Pt 4)*, 913–933. https://doi.org/10.1093/brain/118.4.913

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kayser, J., & Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates [Place: Netherlands Publisher: Elsevier Science]. *Clinical Neurophysiology*, *117*(2), 369–380. https://doi.org/10.1016/j.clinph.2005.08.033

Kelly, S. P., Corbett, E. A., & O'Connell, R. G. (2021). Neurocomputational mechanisms of prior-informed perceptual decision-making in humans. *Nature Human Behaviour*, *5*(4), 467–481. https://doi.org/10.1038/s41562-020-00967-9

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. https://doi.org/10.1523/JNEUROSCI.3355-13.2013

Kim, J.-N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience*, *2*(2), 176–185. https://doi.org/10.1038/5739

Kornhuber, H. H., & Deecke, L. (1965). Changes in the Brain Potential in Voluntary Movements and Passive Movements in Man: Readiness Potential and Reafferent Potentials. *Pflugers Archiv Fur Die Gesamte Physiologie Des Menschen Und Der Tiere*, *284*, 1–17.

Krekelberg, B. (2024). Matlab Toolbox for Bayes Factor Analysis. https://doi.org/10.5281/zenodo.13744717

Lange, F. P. d., Rahnev, D. A., Donner, T. H., & Lau, H. (2013). Prestimulus Oscillatory Activity over Motor Cortex Reflects Perceptual Expectations. *Journal of Neuroscience*, *33*(4), 1400–1410. https://doi.org/10.1523/JNEUROSCI.1094-12.2013

Lenth, R. V., & Piaskowski, J. (2025). Emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 2.0.0. https://github.com/rvlenth/emmeans

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01171

Löffler, A., Haggard, P., & Bode, S. (2020). Decoding Changes of Mind in Voluntary Action - Dynamics of Intentional Choice Representations. *Cerebral Cortex*, *30*(3), 1199–1212. https://doi.org/10.1093/cercor/bhz160

Lui, K. K., Nunez, M. D., Cassidy, J. M., Vandekerckhove, J., Cramer, S. C., & Srinivasan, R. (2021). Timing of Readiness Potentials Reflect a Decision-making

Process in the Human Brain. *Computational Brain & Behavior*, *4*(3), 264–283. https://doi.org/10.1007/s42113-020-00097-5

Maoz, U., Yaffe, G., Koch, C., & Mudrik, L. (2019). Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice (R. G. O'Connell, J. I. Gold, R. G. O'Connell, J. Zhang, & B. Burle, Eds.). *eLife*, *8*, e39787. https://doi.org/10.7554/eLife.39787

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Masaki, H., Takasawa, N., & Yamazaki, K. (1998). Enhanced negative slope of the readiness potential preceding a target force production task. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *108*(4), 390–397. https://doi.org/10.1016/S0168-5597(98)00019-7

McCone, H., Devine, C. A., McNickle, E., Dully, J., Geuzebroek, A. C., McGovern, D. P., Kelly, S. P., & O'Connell, R. G. (2025). A Movement-Independent Signature of Urgency During Human Perceptual Decision Making. https://doi.org/10.1101/2025.05.28.656559

Murphy, P. R., Boonstra, E., & Nieuwenhuis, S. (2016). Global gain modulation generates time-dependent urgency during perceptual choice in humans. *Nature Communications*, *7*(1), 13526. https://doi.org/10.1038/ncomms13526

O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience*, *15*(12), 1729–1735. https://doi.org/10.1038/nn.3248

O'Connell, R. G., & Kelly, S. P. (2021). Neurophysiology of Human Perceptual Decision-Making. *Annual Review of Neuroscience*, *44*(1), 495–516. https://doi.org/10.1146/annurev-neuro-092019-100200

O'Connell, R. G., Parés-Pujolràs, E., Corbett, E. A., Feuerriegel, D., & Kelly, S. P. (2025). Regressing away common neural choice signals does not make them artifacts: Comment on Frömer et al. 2024. *Imaging Neuroscience*, *3*, IMAG.a.60. https://doi.org/10.1162/IMAG.a.60

Ouyang, G., Hildebrandt, A., Schmitz, F., & Herrmann, C. S. (2020). Decomposing alpha and 1/f brain activities reveals their differential associations with cognitive processing speed. *NeuroImage*, *205*, 116304. https://doi.org/10.1016/j.neuroimage.2019.116304

Ouyang, G., Hildebrandt, A., Sommer, W., & Zhou, C. (2017). Exploiting the intra-subject latency variability from single-trial event-related potentials in the P3 time range: A review and comparative evaluation of methods. *Neuroscience and Biobehavioral Reviews*, *75*, 1–21. https://doi.org/10.1016/j.neubiorev.2017.01.023

Ouyang, G., Sommer, W., & Zhou, C. (2015). A toolbox for residue iteration decomposition (RIDE)—A method for the decomposition, reconstruction, and single trial analysis of event related potentials. *Journal of Neuroscience Methods*, *250*, 7–21. https://doi.org/10.1016/j.jneumeth.2014.10.009

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*(11), 1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8

Philiastides, M. G., Heekeren, H. R., & Sajda, P. (2014). Human Scalp Potentials Reflect a Mixture of Decision-Related Signals during Perceptual Choices. *Journal of Neuroscience*, *34*(50), 16877–16889. https://doi.org/10.1523/JNEUROSCI.3012-14.2014

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), 15808. https://doi.org/10.1038/ncomms15808

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Roskies, A. L. (2010). How Does Neuroscience Affect Our Conception of Volition? *Annual Review of Neuroscience*, *33*(1), 109–130. https://doi.org/10.1146/annurev-neuro-060909-153151

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304–321. https://doi.org/10.1037/met0000057

Schurger, A. (2018). Specific Relationship between the Shape of the Readiness Potential, Subjective Decision Time, and Waiting Time Predicted by an Accumulator Model with Temporally Autocorrelated Input Noise. *eneuro*, *5*(1), ENEURO.0302–17.2018. https://doi.org/10.1523/ENEURO.0302-17.2018

Schurger, A., Hu, P. B., Pak, J., & Roskies, A. L. (2021). What Is the Readiness Potential? *Trends in cognitive sciences*, *25*(7), 558–570. https://doi.org/10.1016/j.tics.2021.04.001

Schurger, A., Sitt, J. D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences*, *109*(42), E2904–E2913. https://doi.org/10.1073/pnas.1210467109

Seghezzi, S., Haggard, P., Boyle, G., Northoff, G., Barbey, A., & Fregni, F. (2024). Cognitive Neuroscience of Volition and "Free Will". In *The Sage Handbook of Cognitive and Systems Neuroscience*. Sage Publications Ltd. https://doi.org/10.4135/9781529616651

Selen, L. P. J., Shadlen, M. N., & Wolpert, D. M. (2012). Deliberation in the Motor System: Reflex Gains Track Evolving Evidence Leading to a Decision. *Journal of Neuroscience*, *32*(7), 2276–2286. https://doi.org/10.1523/JNEUROSCI.5273-11.2012

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Soon, C. S., He, A. H., Bode, S., & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences*, *110*(15), 6217–6222. https://doi.org/10.1073/pnas.1212218110

Steinemann, N. A., O'Connell, R. G., & Kelly, S. P. (2018). Decisions are expedited through multiple neural adjustments spanning the sensorimotor hierarchy [Publisher: Nature Publishing Group]. *Nature Communications*, *9*(1), 3627. https://doi.org/10.1038/s41467-018-06117-0

Sun, J., Feuerriegel, D., Nightingale, M., & Osth, A. F. (2025). A Parietal Memory Strength Signal Linked to Evidence Accumulation in Recognition Decisions. https://doi.org/10.1101/2025.07.20.665783

Sun, J., Osth, A. F., & Feuerriegel, D. (2024). The late positive event-related potential component is time locked to the decision in recognition memory tasks. *Cortex*, *176*, 194–208. https://doi.org/10.1016/j.cortex.2024.04.017

Tagliabue, C. F., Veniero, D., Benwell, C. S. Y., Cecere, R., Savazzi, S., & Thut, G. (2019). The EEG signature of sensory evidence accumulation during decision formation closely tracks subjective perceptual experience. *Scientific Reports*, *9*(1), 4949. https://doi.org/10.1038/s41598-019-41024-4

Travers, E., Friedemann, M., & Haggard, P. (2021). The Readiness Potential reflects planning-based expectation, not uncertainty, in the timing of action. *Cognitive Neuroscience*, *12*(1), 14–27. https://doi.org/10.1080/17588928.2020.1824176

Travers, E., & Haggard, P. (2021). The Readiness Potential reflects the internal source of action, rather than decision uncertainty. *European Journal of Neuroscience*, *53*(5), 1533–1544. https://doi.org/10.1111/ejn.15063

Twomey, D. M., Kelly, S. P., & O'Connell, R. G. (2016). Abstract and Effector-Selective Decision Signals Exhibit Qualitatively Distinct Dynamics before Delayed Perceptual Reports. *The Journal of Neuroscience*, *36*(28), 7346–7352. https://doi.org/10.1523/JNEUROSCI.4162-15.2016

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. https://doi.org/10.1111/ejn.12936

van Ede, F., & Nobre, A. C. (2024). A Neural Decision Signal during Internal Sampling from Working Memory in Humans. *The Journal of Neuroscience*, *44*(19), e1475232024. https://doi.org/10.1523/JNEUROSCI.1475-23.2024

van Vugt, M. K., Beulen, M. A., & Taatgen, N. A. (2019). Relation between centro-parietal positivity and diffusion model parameters in both perceptual and memory-based decision making. *Brain Research*, *1715*, 1–12. https://doi.org/10.1016/j.brainres.2019.03.008

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00131

Wyszecki, G., & Stiles, W. S. (1982). *Color Science: Concepts and Methods, Quantitative Data and Formulae. Second Edition* (2nd ed.). John Wiley & Sons, Inc.

Supplementary material for:

Tracing the neural trajectories of evidence accumulation and motor preparation processes during voluntary decisions

Lauren C. Fong[1*], Paul M. Garrett[1], Philip L. Smith[1], Robert Hester[1], Stefan Bode[2] and Daniel Feuerriegel[1]

[1] Melbourne School of Psychological Sciences, The University of Melbourne, Australia

The University of Melbourne, 17 Spencer Road, Parkville VIC 3010, Australia

[2] Division of Science, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Saadiyat Marina District - Abu Dhabi - United Arab Emirates

*Correspondence: fonglc@unimelb.edu.au

Mixed Effects Regression Model Structures and Coefficients for Analyses of Pre-Response Amplitudes, Slopes, and Response Times

As described in the EEG Analysis section of the Methods, we fitted linear mixed-effects models (LMMs) using the lme4 package (Bates et al., 2015) to relate response times (RTs) to pre-response amplitude and slope measures for each signal – the centro-parietal positivity (CPP), Mu/Beta (MB) and left hemisphere readiness potential (LHRP). Models included participant intercepts and, where possible, random slopes for RT, allowing the strength of the RT–measure relationship to vary across participants.

We compared models with and without the fixed effect of RT using maximum likelihood and Chi-Square likelihood-ratio tests within matched random-effects structures as follows:

- m4: $ERP_{measure} \sim 1 + RT + (1 + RT \mid pID)$ vs m3: $ERP_{measure} \sim 1 + (1 + RT \mid pID)$
- m2: $ERP_{measure} \sim 1 + RT + (1 \mid pID)$ vs m1: $ERP_{measure} \sim 1 + (1 \mid pID)$

If the models with random-slope structure (m3 or m4) failed to converge or produced singular fits (i.e., near-zero variance), we defaulted to the models with random-intercept structure (m1 or m2). The best model was chosen as the most complex converged, non-singular model.

Supplementary Table 1

Chi-Square Likelihood Ratio Tests for Group-Level Model Comparisons.

Signal	Decision Type	ERP Measure	Chi-square	df	p	Comparison Test	Best Model
CPP	Voluntary	Slope	20.52	1	<.001***	m1 vs m2	m4
			13.74	1	<.001***	m3 vs m4	
		Amplitude	4.66	1	.031*	m1 vs m2	m3
			1.30	1	.254	m3 vs m4	
	Forced	Slope	18.06	1	<.001***	m1 vs m2	m4
			10.11	1	.002**	m3 vs m4	
		Amplitude	2.56	1	.109	m1 vs m2	m1
			1.44	1	.230	m3 vs m4	
MB	Voluntary	Slope	2.16	1	.141	m1 vs m2	m1
			1.98	1	.160	m3 vs m4	
		Amplitude	0.20	1	.653	m1 vs m2	m1
			0.19	1	.664	m3 vs m4	
	Forced	Slope	0.01	1	.909	m1 vs m2	m1
			0.02	1	.893	m3 vs m4	
		Amplitude	0.51	1	.475	m1 vs m2	m3
			0.56	1	.452	m3 vs m4	
LHRP	Voluntary	Slope	7.23	1	.007**	m1 vs m2	m3
			5.33	1	.021*	m3 vs m4	
		Amplitude	0.32	1	.571	m1 vs m2	m1
			0.10	1	.754	m3 vs m4	
	Forced	Slope	1.55	1	.213	m1 vs m2	m3
			1.06	1	.303	m3 vs m4	
		Amplitude	14.35	1	<.001***	m1 vs m2	m4
			6.44	1	.011*	m3 vs m4	

Note. Models denoted in red produced a singular fit and failed to converge.

*p <.05 **p <.01 ***p <.001

Supplementary Table 2

Linear Mixed-Effects Model Coefficients for Predicting CPP Pre-Response Amplitudes and Slopes from Reaction Times

Decision Type	CPP measure	Model	Estimate	SE	t	p
Voluntary	Amplitude	m2	1.78	0.82	2.16	.031*
		m4	1.55	1.34	1.15	.255
	Slope	m2	-0.02	<.01	-4.53	<.001***
		m4	-0.02	<.01	-4.03	<.001***
Forced	Amplitude	m2	1.20	0.75	1.60	.109
		m4	0.72	1.07	0.68	.502
	Slope	m2	-0.02	<.01	-4.25	<.001***
		m4	-0.02	<.01	-3.34	.002**

Note. m2 denotes the random intercept model. m4 denotes the random slopes model. Rows in red indicate results from models that failed to converge or produced a singular fit. *p <.05 **p <.01 ***p <.001

Supplementary Table 3

Linear Mixed-Effects Model Coefficients for Predicting MB Pre-Response Amplitudes and Slopes from Reaction Times

Decision Type	MB measure	Model	Estimate	SE	t	p
Voluntary	Amplitude	m2	0.03	0.07	0.45	.031*
		m4	0.03	0.07	0.43	.255
	Slope	m2	<.01	<.01	1.47	<.001***
		m4	<.01	<.01	1.35	<.001***
Forced	Amplitude	m2	-0.05	0.07	-0.71	.109
		m4	-0.06	0.09	-0.75	.502
	Slope	m2	<.01	<.01	0.11	<.001***
		m4	<.01	<.01	0.14	.002**

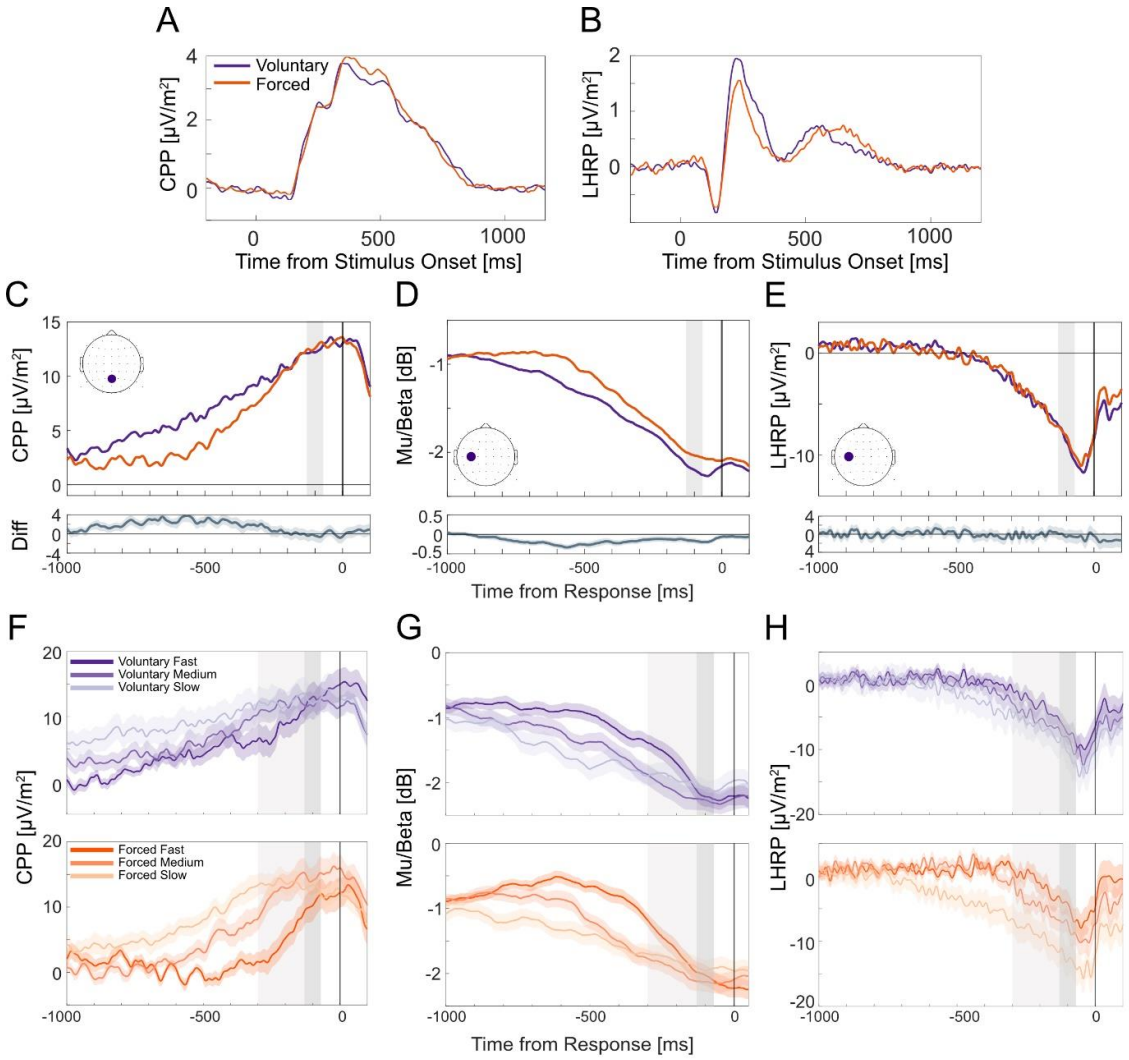
Note. m2 denotes the random intercept model. m4 denotes the random slopes model. Rows in red indicate results from models that failed to converge or produced a singular fit. *p <.05 **p <.01 ***p <.001

Supplementary Table 4

Linear Mixed-Effects Model Coefficients for Predicting LHRP Pre-Response Amplitudes and Slopes from Reaction Times

Decision Type	LHRP measure	Model	Estimate	SE	t	p
Voluntary	Amplitude	m2	-0.56	0.99	-0.57	.571
		m4	-0.36	1.16	-0.31	.754
	Slope	m2	0.01	<.01	2.69	.007**
		m4	0.01	<.01	2.42	.019*
Forced	Amplitude	m2	-3.44	0.91	-3.79	<.001***
		m4	-3.53	1.34	-2.62	.012
	Slope	m2	-0.01	<.01	-1.25	.213
		m4	-0.01	0.01	-1.03	.306

Note. m2 denotes the random intercept model. m4 denotes the random slopes model. Rows in red indicate results from models that failed to converge or produced a singular fit. *p <.05 **p <.01 ***p <.001



Supplementary Figure S1. CPP, Mu/Beta, and LHRP amplitude waveforms. A-B) RIDE-derived S (stimulus-locked) subcomponent group-average waveforms at electrode Pz (A) and C3 (B), reflecting ERP waveform contributions that are strictly locked to stimulus onset. C) Group-averaged CPP, D) MB amplitudes, and E) LHRP for voluntary (purple) and forced (orange) waveforms with accompanying difference waves (and shaded standard error regions) at each time point. Plotted waveforms are for data where the S component was not subtracted and the overlap of stimulus-locked waveforms was not corrected for. Grey areas represent the pre-response windows between -130 to -70 ms. F) Group averaged CPP, G) MB amplitudes, and H) LHRP for voluntary (purple) and forced (orange) decisions split by fast, medium, and slow RT bins, with shading denoting standard errors. Dark grey areas denote the -130 to -70 ms pre-response amplitude window. Light grey areas denote the -300 to -70 ms time slope window. Overall, the waveforms are very similar to the data where the S component had been subtracted.

Supplementary Material References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.