How spatial attention affects the decision process: looking through the lens of Bayesian hierarchical diffusion model & EEG analysis

Ghaderi-Kangavari, A.; Parand, K.; Ebrahimpour, R.; Nunez, M.D.; Rad, J.A.

DOI
10.1080/20445911.2023.2187714

Publication date
2023

Document Version
Final published version

Published in
Journal of Cognitive Psychology

JOURNAL OF COGNITIVE PSYCHOLOGY
2023, VOL. 35, NO. 4, 456–479
https://doi.org/10.1080/20445911.2023.2187714

How spatial attention affects the decision process: looking through the lens of Bayesian hierarchical diffusion model & EEG analysis

Amin Ghaderi-Kangavari [a], Kourosh Parand [a,b], Reza Ebrahimpour [c,d,e], Michael D. Nunez [f] and Jamal Amani Rad [a]

[a]Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran; [b]Department of Data and Computer Sciences, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran; [c]Center for Cognitive Science, Institute for Convergence Science and Technology, Sharif University of Technology, Tehran, Iran; [d]Faculty of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran; [e]School of Cognitive Sciences (SCS), Institute for Research in Fundamental Sciences (IPM), Tehran, Iran; [f]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

ABSTRACT

We explored the underlying latent process of spatial prioritisation in perceptual decision processes, based on the drift-diffusion model, and subsequent nested model comparison. Our hierarchical cognitive modelling analysis revealed that spatial attention changed the non-decision time parameter across experimental conditions, quantified using the deviance information criterion score (DIC) and R-squared. We also constructed joint models with embedded neural covariates to discover which of contralateral and ipsilateral EEG measures could most manipulate spatial attention in perceptual decision making. Using multiple regression analysis, it can be concluded that poststimulus N2nc can predict mean response times and non-decision time parameters related to spatial prioritisation. However, the contralateral minus neutral alpha power at parieto-occipital electrodes can only predict the mean RTs and not the non-decision time relating to spatial prioritisation. Results suggest that individual differences in spatial attention are encoded more by contralateral (and not ipsilateral) N2 oscillations and non-decision times.

ARTICLE HISTORY
Received 8 April 2022
Accepted 7 December 2022

KEYWORDS
Perceptual decision making; visual spatial attention; diffusion decision model; EEG; hierarchical Bayesian inference

1. Introduction

Decision-making is a high-level cognitive process whereby the decision-maker makes decisions based on the available evidence, expected value, and the possible outcomes (Gold & Shadlen, 2007). The decision process is assumed to accumulate noisy evidence in a flexible time frame, instead of focusing on immediate evidence acquisition (Gherman & Philiastides, 2018; Shadlen & Kiani, 2013). The most frequent type of decision that we routinely make in different situations is perceptual decision making, a process that involves the processing of information, the accumulation of evidence, and motor response (Philiastides et al., 2006). One of the independent variables that plays a key role in the perceptual decision-making process is spatial attention (Gluth et al., 2018; Klatt et al., 2020; Krajbich et al., 2012; Nunez et al., 2017). For instance, suppose that you are driving to an intersection on a sunny day with the help of traffic lights; it is easy to determine whether to continue driving or stop the vehicle. But if the situation changes and sensory information is disturbed, then response time and accuracy are affected. One example is that when it is rainy and dark (i.e. a decrease in visual coherence) a driver will not effortlessly detect the colour and symbol of the traffic light, and thus their ability to stop or accelerate could be affected. A similar, but different effect, is when a driver is moving toward the intersection in rainy and dark weather, and she/he does not know the exact spatial location of the traffic light. The driver could initially struggle to locate the traffic light (i.e. the stimulus) before processing the received information from the light. If the driver is told exactly where the light is (e.g. spatial cueing) before reaching the intersection, the driver could focus their attention on the target and suppress

CONTACT Jamal Amani Rad j amanirad@sbu.ac.ir; j.amanirad@gmail.com Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran

Supplemental data for this article can be accessed http://doi.org/10.1080/20445911.2023.2187714.

attention to other areas of visual space faster and maybe more acmadulationcurately (Nunez et al., 2015; Ostwald et al., 2012; Shadlen & Kiani, 2013; Yeshurun & Carrasco, 1999). Posner's (Posner, 1980) studies have shown that people can pay close attention to an area of space without moving their eyes to that area. Moreover, one type of spatial attention is goal-driven, suggesting top-down or endogenous attention in situations where we attend to a region of space where the stimulus is present (i.e. valid) or not present (i.e. invalid) (Posner, 1980). The top-down cue could facilitate behavioural performance and manipulate neural mechanisms (Imani et al., 2021; Sagar et al., 2019), of which some aspects are still unknown. In this paper, we focus on valid, covert spatial attention to stimuli which were presented by endogenous orienting cues. The role of spatial attention in perceptual decision-making has not yet been identified by emphasising modelling and neural measures simultaneously, so the present research is an attempt to fill this gap.

One measure with a rich history in the study of spatial attention is the alpha band (8-13 Hz) given by frequency and time–frequency compositions of electroencephalogram (EEG). Top-down spatial attention is thought to modulate the alpha band (8-13 Hz) of downstream areas such as posterior-occipital sites to direct and allocate limited processing resources to attended and unattended loci (Foster et al., 2017; Li et al., 2013). In fact, studies show that increasing alpha band power ipsilaterally (alpha-synchronization) to the unattended location could suppress task-irrelevant space, and the decreasing alpha band power contralaterally (alpha-desynchronisation) to the attended location intentionally facilitates upcoming visual processing (Foster et al., 2017; Rihs et al., 2007). This mechanism of top-down orienting of attention is usually found after post-cue (pre-stimulus) of various tasks such as visual (Praamstra et al., 2005; van Schouwenburg et al., 2017), tactile (Haegens et al., 2012), and auditory (Bernier et al., 2017) tasks.

Another class of EEG measure that is of interest in spatial attention is the negative deflections in trial-averaged ERPs after visual stimuli occur, in particular the second negative peak (N2) amplitudes. N2s have multiple interpretations and differences in amplitude and latency depending upon the task and location over electrodes. One example is the central contralateral N2 (N2cc) subcomponent in central electrode sites. The N2cc has been shown to have a substantial role in the prediction of distributed finite resources for spatial prioritisation (Praamstra, 2006). The N2cc amplitude, which is preparation thought to mirror a motor mechanism, decreases with informative top-down cues (Praamstra, 2006). The N2cc may also reflect an attention-related motor component which could prevent the release of incorrect responses after the presentation of stimuli in the ipsilateral field (Amenedo et al., 2012). Moreover, researchers have reported that the N2cc over central electrodes sites might coincide in time with the posterior contralateral N2 (N2pc), but that the important finding of N2cc is not due to overlap of the N2pc with movement execution-related activity (Praamstra, 2006; Praamstra & Oostenveld, 2003). It has also been shown that the age-related behaviour was related to N2pc and N2cc subcomponent during visual search task (Amenedo et al., 2012). The N2pc was found to play a role in visuospatial processing of target stimuli and the control of non-target stimuli, and the N2cc was also correlated to inhibition of spatial response tendencies in the Simon task (Cespon et al., 2016). Another similar measure, the anterior contralateral N2 subcomponent (N2ac) has been shown to predict sound localisation performance and diffusion model parameters using multiple linear regression (Gamble & Luck, 2011; Klatt et al., 2020). In particular, it has been shown that N2ac amplitudes modulated accuracies and evidence accumulation rates estimated by drift-diffusion models (DDMs) (Klatt et al., 2020). Finally, generalised contralateral (N2c) and ipsilateral (N2i) N2 subcomponent peak amplitudes at electrodes P7/P8 have been found to change as a function of motion strength (Loughnane et al., 2016), and generalised N2 amplitudes have been shown to track non-decision times in DDMs (Nunez et al., 2019).

EEG signals thus contain informative signals encoding spatial attention that describe the allocation of resources to speed up perceptual decision making. However, in the literature it is unclear if these EEG signals are contralateral or ipsilateral to the attended stimulus. It could be that some contralateral and ipsilateral EEG measures of spatial attention do not describe anything about spatial attention, and this would not be clear from the literature since most studies calculate EEG biomakers from the difference in contralateral and ipsilateral signals (Cespon et al., 2016; Gamble & Luck, 2011; Klatt et al., 2020; Luck, 2014). Therefore, in our study, we purposely compared contralateral and

ipsilateral EEG signals to a neural condition, to better understand how EEG measures describe cognition through direct testing of cognitive variables and behaviour. Imani et. Al (Imani et al., 2021) used the current dataset to determine neural signature of the information processing stage by employing the hidden semi-Markov model method (HSMM). They employed a new approache to fill a gap between physiological data and behavioural components. Then, they used some preprocessed snapshots to bring down dimensionality and then fed them to the HSMM-EEG to split up the general three stages in perceptual decision-making; encoding, decision and response. As a result, the role of contralateral and ipsilateral in spatial attention for the data are still unknown, and also model selection and nested model comparisons are more helpful to identify the most important role of spatial attention in the DDM's parameters which are not implemented to the 2×2 factorial data.

In general, SSMs assume that decisions are made from a noisy process of accumulating evidence where the evidence is gradually accumulated over time until a sufficient amount of evidence for one of the choices reaches a predetermined threshold, at which point a decision is made. DDMs assume particular random processes for accumulating evidence, "Wiener" or Brownian motion-based processes. The reason for this assumption is that evidence accumulation is thought to be a process in which the individual may respond differently to the same stimulus in different trials (Choi & Paik, 2019; Ratcliff et al., 2016). DDMs can also be considered as a bridge between behavioural data and cognitive processes because DDMs have cognitive parameters that have been experimentally-supported interpretations (Ratcliff et al., 2016; Voss et al., 2004). The purpose of behavioural modelling is to understand what components or parameters can be extracted from the perceptual decision making process, how different people make correct responses and commit errors, and the time course of these processes (Forstmann et al., 2016). There is a growing consensus that the evidence accumulates gradually and sequentially to make a decision (Ratcliff, 1978; Ratcliff & McKoon, 2004; Smith, 2016). As a result, sequential sampling models have become the most well-known explanation of how the decision-making process works (SSMs; see Stone, 1960 (Stone, 1960), Ratcliff et al., 2016 (Ratcliff et al., 2016) and Evans & Wagenmakers, 2020 (Evans & Wagenmakers, 2020) for reviews). These models have been successful in describing the cognitive processes underlying decision making across a wide variety of paradigms, such as studies of optimality (Drugowitsch et al., 2012; Evans et al., 2019; Evans et al., 2020; Evans & Brown, 2017), stop signal paradigms (Matzke et al., 2013), go/no-go paradigms (Gomez et al., 2007), multi-attribute & many alternatives choice (Kvam, 2019; Mallahi-Karai & Diederich, 2019; Usher & McClelland, 2004), learning strategies (Fontanesi et al., 2019; Pedersen et al., 2017; Sewell et al., 2019), attentional choice (Gluth et al., 2018, 2020; Krajbich et al., 2012; Nunez et al., 2017), continuous responses (Ratcliff, 2018; Smith, 2016), neural processes (Gold & Shadlen, 2007), and so on.

Neurocognitive models are powerful tools to link EEG to underlying latent cognitive model parameters in attention-related tasks (Nunez et al., 2017; Nunez et al., 2019; Palestro et al., 2018). In the directed neurocognitive models, neural covariates constrain and predict DDM parameters directly (Frank, 2015; Nunez et al., 2019; Turner, Forstmann, et al., 2019a). This approach is easy to understand and implement, analogous to multiple regressions between cognitive latent variables and neural measures. However, these models are fit in a single step in a Bayesian framework in order to capture full uncertainty about the effects of EEG influence on model parameters. We also created directed neurocognitive models in a *hierarchical* Bayesian framework to estimate posterior parameters at individual and group levels (Lee, 2011).

In this paper, we explore which DDM parameters are manipulated by spatial attention. By nested model selection criteria, we show that non-decision time is integral to encoding spatial prioritisation. We also found the measures of spatial prioritisation in the electroencephalogram (EEG) to predict behavioural performance and DDM parameters in a face-car perceptual decision making task with spatial cueing manipulations. To be more specific, we tested whether the lateralised alpha frequency band at parieto-occipital sites together with the contralateral-neutral (contralateral minus neutral difference curve) and ipsilateral-neutral (ipsilateral minus neutral difference curve) N2 subcomponents at central sites could modulate behavioural performance and parameters of spatial prioritisation in perceptual decision making (Klatt et al., 2020; Praamstra & Oostenveld, 2003). The goal of this study was to find the relationship between these two intriguing EEG measures, e.g. the lateralised

alpha frequency band amplitudes over parieto-occipital electrode sites and N2-subcomponent amplitudes, with behavioural performance and the non-accumulation time, also called the non-decision time, as estimated from DDMs (Ratcliff & McKoon, 2004). We applied hierarchical Bayesian modelling with a powerful nested model comparisons approach to estimate DDM model parameters and find the model that best described the data (Olianezhad et al., 2019). Then we found out how well individual differences in EEG measures of spatial prioritisation described individual differences in task behaviour and DDM parameters.

In summary, we hypothesise that electrophysiological correlates of top-down spatial prioritisation (alpha band lateralisation and N2 subcomponents) during a visual perceptual decision making task could predict visual performance and the non-decision time parameter of DDMs, reflecting the sensory coding time and response expectation. Our claim is that the cognitive process of spatial attention is encoded by lateralised alpha power in and around parieto-occipital electrode sites, and that spatial attention is also encoded by lateralised N2cc amplitudes in central electrode sites. These measures of spatial attention in EEG should regulate the speed to choose a face and car stimuli, so they should contribute to the non-decision processes and the performance.

2. Methods

2.1. Experiment overview

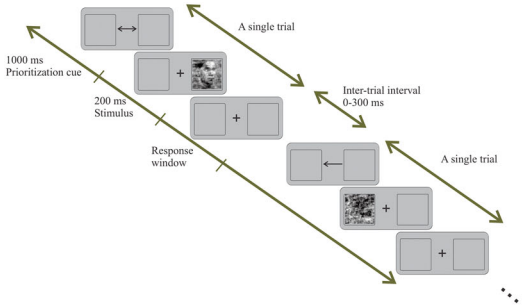
In this study, we re-examined the data from an experiment conducted by (Georgie et al., 2018) to understand visual coherence and spatial attention effects in perceptual decision-making. In their study, seventeen participants (8 females, mean age was 25.9 years, range 20–33 years, 2 left-handed) from the University of Birmingham completed the experiment, which was approved by the University of Birmingham Ethical Review Committee, and were paid 7.5 per hour with travel expenses reimbursed. Electroencephalograms (EEG) and behavioural data from a perceptual decision making task with spatial cueing were acquired over two data acquisition sessions.

As shown in Figures 1a and b, on each experimental trial, in order to manifest the informative and uninformative spatial prioritisation, cueing one-way arrows or two-way arrows were shown for a one-second period before the stimulus emergence. Then, a visual stimulus of a car or face was displayed on the screen for 200 milliseconds. Participants then pressed a button to respond whether they perceived the car or the face. Participants used their index finger and middle finger of the right-hand to respond. This was followed by an inter-trial period between 0 msec and 300 ms. It should be noted that participants were also told to maintain fixated on the fixation point during the entire trial.

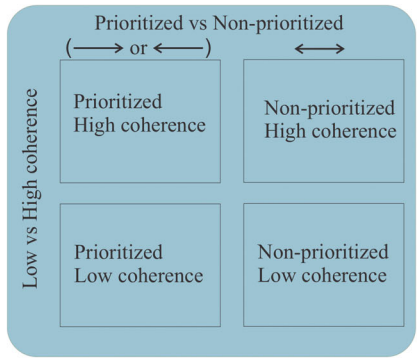
The experimental manipulation of coherence had two levels: “low” and “high”, and experimental manipulation of spatial cueing had also two levels: “yes” and “no”, such that a leftward or rightward facing arrow was presented in half of the trials (“yes”) and a double-sided arrow was presented in half of the trials (“no”). These arrows indicated to participants that they should attend to one side (informative cueing) of space or both sides of space (uninformative cueing) covertly. All trials from each experimental condition and experimental stimuli (i.e. face or car) were randomly presented to participants. For more details about experimental design & stimuli, interested readers can refer to the main report written on the data by Georgie et al. (Georgie et al., 2018).

Following (Georgie et al., 2018), 36 trials of both EEG and behavioural data were acquired for each of the experimental conditions (coherence and spatial attention) and two stimulus types (face or car). The EEG was collected with a sampling rate of 5000 Hz samples using 64 EEG sensors, which consisted of 62 scalp electrodes via the 10–20 system and two extra sensors. One extra sensor was placed approximately 2 cm under the left collarbone in an effort to identify electrocardiogram (ECG) signals, and the other sensor was placed under the left eye to aid in correcting eye-blinking artifacts (EOG).

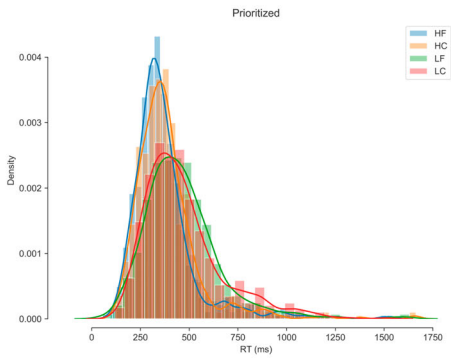
In our study, three participants were excluded because their EEG data were too contaminated with the artifact, which left fourteen participants in our analysis. Also, in order to perform statistical analysis and model fitting, we removed outliers in the behavioural data. We used an interquartile range (IQR), with IQR = Q3 - Q1 as the distance from the first quartile to the third quartile of response times. To compute the interquartile range for the data, any response time smaller than



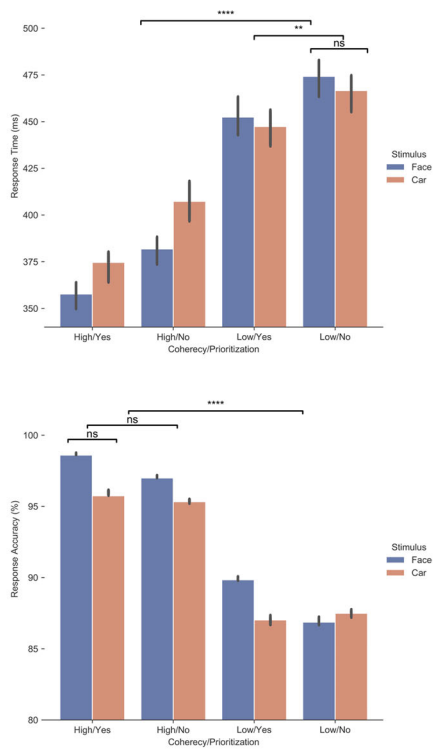
(a) For each trial, one-way arrows (informative cueing) or two-way arrows (uninformative cueing) were shown in the center of the screen for 1000 milliseconds. Then, a picture of a car or face was shown on left or right side of the screen for 200 milliseconds. Participants then pressed a button to respond whether they perceived a car or a face.



(b) The study design was a 2×2 factorial design. Twolevels were given to participants for coherence (highand low) and two levels for spatial attention (prioritized vs non-prioritized), randomly mixed across trials.



(c) Response time (RT) distributions collapsed across participants for cued (Prioritized) and uncued (Non-prioritized) stimuli. There are four overlapping RT distributions in each of the two subplots for high coherence trials with face stimuli (HF) and car stimuli (HC). As well as low coherence trials with face stimuli (LF) and car stimuli (LC).



(d) Mean Response Times and Mean Accuracies averaged over participants with standard error bars for each experimental condition (High or Low coherence; Yes or No Spatial Cueing) and for each stimulus (Face or Car). Spatial cueing and coherence have significant effects on response time, but stimulus did not significantly manipulate response time. Only coherence affected accuracy across individuals.

Figure 1. Experimental design and behavioural data. (A-B) Overview of the experiment task and the 2 × 2 study design. (C-D) Bar plots and error bars displayed the grand-average and ±1 standard error of the mean (SEM) response time across independent variables. Asterisks indicate significant *p*-value as $^{**}p < 0.01$ and $^{****}p < 0.0001$, and "ns" indicates non-significant *p*-value. [To view this figure in color, please see the online version of this journal.]

$Q3 - 3 \times IQR$ and greater than $Q3 + 3 \times IQR$ were removed as outliers resulting in 53 trials being removed. Consequently, response time and accuracy are submitted to parametric $2 \times 2 \times 2$ repeated-measure analysis of variance (ANOVA) statistical test.

2.2. EEG signal analysis

The EEG preprocessing and analysis were implemented by the MNE-python package (Gramfort et al., 2013). The preprocessing steps we performed can be summarised as follows: the continuous EEG data were down-sampled to 256 samples, the EEG data were band-pass filtered to the range of 1 Hz - 30 Hz, EEG was re-referencing to the common average, visual inspection of the EEG was performed to remove trials with high noise amplitude or abnormal activity, signals containing too much noise were interpolated, trials with amplitudes outside the range of −100 mV to +100 mV were removed, EEG data was split into epochs from −100 msec (baseline) time-locked to the cue's onset to 600 msec after stimulus appearance (totalling 1700 ms), Independent Component Analysis (ICA) was run to manually remove EEG data irrelevant to the task (e.g. eye movement, head motion and muscular activity), Independent Components (ICs) were removed automatically that best matched EOG sensors or ECG sensors (such that these ICs reflected eye blinks and heart rhythm artifacts respectively) using the MNE-package (Gramfort et al., 2013), and finally the EEG data was converted back into sensor space from IC space. The EEG preprocessing with MNE-package implementation code in python named "MNE-preprocessing" is readily accessible at https://github.com/AGhaderi/MNE- Preprocessing.

In order to explore the spectral content of the data, we employed a Morlet wavelet transform to the EEG in Python. The Morlet wavelet is an efficient time–frequency decomposition over every trial. We used this time–frequency approach in order to appropriately understand the properties of the non-stationary EEG signal, and to keep the random phase information before averaging across trials (in comparison to the event-related potential analysis discussed below, which only extracts phase-locked information across trials). We changed the parameter of "number of cycles", which controls the trade-off between temporal precision and frequency precision in this algorithm, by increasing by 0.5 cycles per frequency, starting at 4 Hz, in steps of 1 Hz. That is, we used 3-cycle wavelets at the lowest frequency (4 Hz), 3.5-cycle wavelets at the next lowest frequency (5 Hz), and so on until we reached 16-cycle wavelets at the highest frequency (30 Hz). After performing this single-trial Wavelet decomposition from 4Hz to 30Hz, the time–frequency decomposition per trial was averaged across trials for each participant. We used this time–frequency average across trials to extract the alpha band from 8 to 13 Hz per participant by averaging the power from 8 to 13 Hz.

We also employed an event-related potential (ERP) analysis. This analysis better captures the phase-locked information across trials (as opposed to the time–frequency analysis discussed previously). Because EEG data routinely contains both signals from real brain sources but also from biological or non-biological artifact and noise even after rigorous artifact correction (Luck, 2014; Nunez et al., 2016), a phase-locked analysis is useful to better average out or eliminate outlier artifact. Therefore, we used an ERP approach to detect cognitive components across experimental conditions and for each participant. We used a baseline of 100 msec before the cue onset in order to calculate the N2 amplitudes after the stimulus appearance.

Both the N2 ERP subcomponent waveforms and alpha band frequency amplitudes were viewed by eye before being used in regression models. We calculated the mean contralateral and ipsilateral (informative) and neutral evoked potentials (uninformative) at C1/2/3/4 central electrodes, and also the mean ipsilateral and contralateral and neutral power across time in the alpha frequency band (8–13 Hz) at parieto-occipital electrodes PO3/4/7/8 (Amenedo et al., 2012; Praamstra, 2006; Praamstra & Oostenveld, 2003). We calculated the contralateral responses as the average of ERP oscillations in left hemispheric sensors during right-arrow cue trials and in right hemispheric sensors during left-arrow cue trials. Similarly, we calculated the ipsilateral responses as the average ERP oscillations at left hemispheric sensors in left-arrow cue trials and right hemispheric sensors in right-arrow trials. We also calculated the neutral responses as the average of ERP oscillations at both left and right hemispheric sites in two-way arrow trials.

Many spatial attention studies have focused on pre-stimulus components and alpha oscillations (Sagar et al., 2019; Wilsch et al., 2020), and more specifically previous researchers have focused on pre-stimulus alpha oscillations as the anticipatory indexes for allocating spatial prioritisation (Van Dijk et al., 2008). However, we concentrated on post-stimulus components during perceptual decision making tasks to target the actual deployment of spatial attention (Klatt et al., 2020). In order to measure the N2 subcomponent

amplitudes, we computed the dependent t-test statistic over contralateral (at C1 & C3 or C2 & C4 in informative cue conditions) minus neutral (at C1, C2, C3, and C4 in uninformative cue conditions) amplitude (N2nc) and ipsilateral minus neutral amplitude (N2ni) at a time window within 150–500 msec after stimulus onset (1150-1500 msec after cue onset). This resulted in 90 statistical t-tests, with a test for each time point in the 350 msec window with a 256 sample rate, so we used a False Discovery Rate (FDR) correction procedure that gave us a different standard for significance to correct the t-tests (Luck, 2014). The measurement window of N2nc and N2ni amplitudes were then based on significant time samples by FDR correction. We then used the variables of {N2nc}, and {N2ni} were used as regressors in the multiple regression model described later.

To assess alpha lateralisation, two contralateral and ipsilateral alpha band indices were used across both hemispheres as follows:

Anc = contralateral alpha power (at PO3 & PO7 or PO4 & PO8 in informative cue conditions) minus neutral alpha power (at PO3, PO4, PO7, and PO8 in uninformative cue conditions)

Ani = ipsilateral alpha power (at PO3 & PO7 or PO4 & PO8 in informative cue conditions) minus neutral alpha power (at PO3, PO4, PO7, and PO8 in uninformative cue conditions)

When the Anc index is negative, the contralateral alpha power is lower than the neutral alpha power at parieto-occipital sensors and vice versa. In contrast, when the Ani index is positive, the ipsilateral alpha power is higher than the neutral alpha power and vice versa. To determine the measurement window of alpha power in the grand-averaged Anc and Ani indices, we used the 50% Fractional Area Latency (FAL). The FAL estimates the midpoint latency of a component signal (Klatt et al., 2020; Luck, 2014). We calculated 50% FAL contingent on a wide time window, from −100 to 600 msec after stimulus onset (i.e. 900–1600 msec after cue onset). We then used a 200 msec time window around the FAL to extract relevant Anc and Ani. The variables of {Anc}, and {Ani} were then used as regressors in the multiple regression model in addition to the {N2nc}, and {N2ni} regressors mentioned previously.

2.3. Drift-diffusion model

Hitherto, various sequential sampling models have been presented by researchers. There are some successful sequential sampling models such as: EZ-diffusion model (Wagenmakers et al., 2007), linear ballistic accumulators (Brown & Heathcote, 2005), leaky-competing accumulator (Usher & McClelland, 2001), and race diffusion model (Tillman et al., 2020), but the drift-diffusion model (DDM) is the most popular model of evidence accumulation (Ratcliff, 1978). Similar to the other sequential sampling models, the DDM is based on evidence accumulation until reaching a threshold. In other words, DDM assumes that the information accumulation process starts from a point between two fixed boundaries, the accumulator steps toward the upper/lower boundary, and after the accumulator passes one of these boundaries, the accumulation process stops and the corresponding option is selected (Ratcliff & Smith, 2015), illustrated in Figure 2. Mathematically, evidence accumulation in the DDM can be described continuously as,.

$$dX_t = v \cdot dt + dW_t\ , \tag{1}$$

or approximated discretely as,.

$$X(t + \Delta t) = X(t) + v \cdot \Delta t + e \cdot \sqrt{\Delta t}, \tag{2}$$

in which X_t is the diffusion state and dW_t denotes the standard Wiener process. Also, in discrete form (simulated in Figure 2), e is normally distributed noise $N(0, 1)$. Traditionally, the DDM consists of four major parameters. The first parameter is the threshold. In DDM the lower boundary is at zero and the upper boundary has a distance "a" from the lower boundary. Therefore, "a" is the threshold parameter, and actually is an index for the speed-accuracy tradeoff. That implies that for higher values of "a", the participant accumulates more evidence to make a decision, and thus his/her decisions will be accurate more often, but often slower. On the other hand, lower values of "a" imply that the participant often makes the decision faster but will be accurate less often (Evans & Brown, 2017). The second parameter of the DDM is the drift rate, which is denoted by "v". The drift rate is the mean slope of each step of the accumulator toward the boundaries. This parameter encodes the speed of information processing and identifies task difficulty levels. Higher values of the drift rate show that the task has no demanding cognitive load and that it is easy to make a decision between the options. When the value of the drift rate tends to zero, this implies the task is demanding and the speed of information processing is

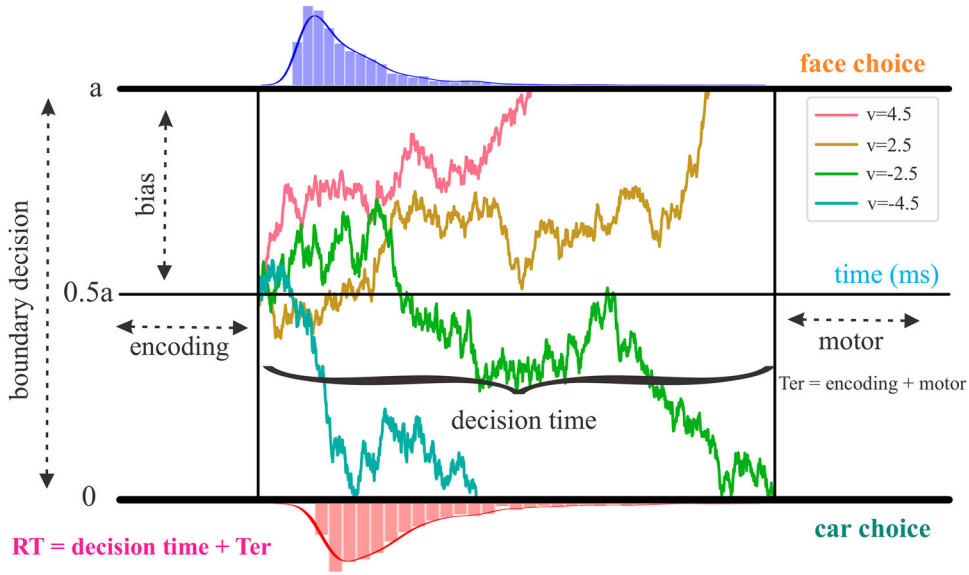


Figure 2. Trajectories of drift-diffusion model for two stimuli with various drift rates. Four trajectories of evidence accumulation are shown on 4 trials with a different drift-rate *v* in each trial. Pink, yellow, green, and blue trajectories were generated with drift rates of 4.5, 2.5, −2.5, and −4.5 respectively. When enough evidence is reached by crossing one of two boundaries for a "face" or "car" choice, a decision is made on that trial. The time course of the trajectories encodes the decision time on that trial. The decision time shown in this figure is of the trial given by the green trajectory. "encoding" and "motor" non-accumulation times are assumed to be the two additive components of non-decision time T_{er}. In these four trials, there was no initial bias towards a "face" or "car" choice, such that the start point is equal to 0.5*a*. [To view this figure in color, please see the online version of this journal.]

slow (Voss et al., 2013). The third parameter is the bias parameter, which is denoted by "*z*", that encodes the starting point of the accumulation process. When the starting point is equal to "*a*", the distance between the starting point and the boundaries are equal and there is no bias to each option. But when "*z*" is greater or less than "$a/2$ ", then the participant has a bias to upper or lower boundaries, respectively (Ratcliff & Rouder, 1998). The fourth parameter is the non-accumulation time T_{er}, also called the non-decision time. This parameter is added to the model for the purpose of excluding the encoding and motor time for the response time. Usually, this parameter is denoted by T_{er} in DDM (Voss et al., 2004). These are the main parameters of the DDM, however often some between-trial variability in parameters is assumed. s_v, s_z, and s_t are the between-trial variability parameters for drift rate, bias, and non-decision time respectively. Usually, normal distributions are assumed for *sv* and uniform distributions for *sz* and *st* (Ratcliff & McKoon, 2004). High values for the between trial variability parameters indicate that there is some variability in the stimuli across trials or in cognitive states (Ratcliff & Tuerlinckx, 2002).

2.4. Cognitive modelling and model comparison

We used a hierarchical Bayesian estimation approach in order to explore the latent cognitive processes underlying perceptual decisions during spatial prioritisation, and to link those latent cognitive processes to neural mechanisms. We assumed hierarchical models where the participant-specific parameter values are randomly sampled from a group-level distribution. And then we used a hierarchical Bayesian approach such that the marginal distributions of the parameters are estimated simultaneously at the group level and the individual level (Vandekerckhove et al., 2011) for each model. We combined this hierarchical approach with the diffusion model in a stable estimation of parameters (i.e. drift rate, boundary separation, and non-decision time) at both individual and group levels. Specifically, we fit models using the HDDM package in Python (Wiecki et al., 2013), which generates full posterior distributions of the model parameters at the individual and group level via Markov-chain Monte Carlo sampling (MCMC). We used the HDDM package because is flexible and user-friendly. We also applied hierarchical Bayesian

estimation because it returns stable results and appropriate parameter estimation in the form of posterior distributions for each parameter included in the model, even when less data (i.e. experimental trials) are available compared to other implementations such as traditional maximum likelihood estimation for individual participants (for details about the model see (Ratcliff & Childers, 2015; Wiecki et al., 2013)).

The HDDM package offers different model parameterizations depending on the experimental factors, e.g. one could model different drift rates for all conditions or all groups or any combination of them (Spilcke-Liss et al., 2019). Different criteria such as the deviance information criterion (DIC) can be used for a comparison across these models. In the full DDM, there are seven parameters which could be categorised into three parts: the decision process measures (boundary decision, mean starting point, and mean drift rate), the non-decision process measure (non-decision time), and the inter-trial variability measures (drift rate variability s_v, starting point variability s_z, and non-decision time variability s_t) (Choi & Paik, 2019; Ratcliff et al., 2016). Based on the strong evidence previously presented in various studies to investigate the effects of coherence and stimulus on the decision-making process (Olianezhad et al., 2019; Philiastides et al., 2006; Ratcliff et al., 2009), we first tested cognitive models whose drift rate parameters are dependent both on the coherence condition (high or low) and stimulus (face or car).

In order to test our hypothesis that spatial attention regulated resources of decision making, we compared five possible nested models (Olianezhad et al., 2019). First, the spatial cueing (i.e. directional arrows presented to participants) may not capture or shift any resources or decision parameters, so this model has no parametric dependence on spatial prioritisation. We labelled this model $model_p$, in which only the drift rate parameter depends on coherence and stimulus (face or car). Second, spatial cueing could regulate the rate of accumulated evidence (the drift rate v), labelled $model_v$. Third, non-decision time (T_{er}) could vary with spatial cueing, labelled $model_t$. Four, spatial cueing could manipulate the mean bias (starting point z), labelled $model_z$. Five, spatial cueing might shift the decision boundary (a), called $model_a$. Note that in each model the drift rate v was free to vary with the level of the coherence (high or low), in line with related findings about stimulus strength in perceptual tasks (Philiastides et al., 2014, 2006). Moreover, inter-trial variability in non-decision time $'s_t'$ was a free parameter in all models on the group level, whereas other inter-trial variables were not estimated. Note that there are two ways to implement DDMs: an accuracy-based approach in which correct responses are coded 1 and incorrect responses 0, or a stimulus-based approach in which the response choice, such as face or car, is modelled. In this work, the stimulus-based approach is employed in which "face" responses are coded 1 and "car" responses are coded 0.

For each model, Markov Chain Monte Carlo simulations (Gamerman & Lopes, 2006) were used to generate 100,000 samples from the joint posterior distribution of parameters using the HDDM package (Wiecki et al., 2013), from which the initial 1000 samples were discarded as a burn-in phase to minimise the effect of initial values on the posterior inference and converge on stable distributions (Wiecki et al., 2013). The convergence of the Markov chains was assessed through visual inspection as well as by calculating the Gelman-Rubin statistic (Gelman & Rubin, 1992), R-hat, to ensure that the models had properly converged, such that R-hat compares between-chain and within-chain variance.

Both R-squared (R^2) and the deviance information criterion (DIC) of each model were used to evaluate each model's goodness-of-fit while accounting for model complexity (i.e. number of free parameters), with lower DIC values indicating better model fit (Spiegelhalter et al., 2002). Note that DIC has been previously used to choose the best diffusion model at group level (van Schouwenburg et al., 2017; Wiecki et al., 2013). DIC score is reported for group levels of DDMs. R^2 was generated from simulated RT and choice data from the group level DDM parameters across conditions by comparing generated mean RTs and accuracies with observed group mean RTs and accuracies.

To further evaluate the best fitting model, we ran posterior predictive checks by averaging 500 simulations generated from the model's posterior to confirm it could reliably reproduce patterns in the observed data (Wiecki et al., 2013), see Supplementary Material. The hDDM implementation code and simulation scripts for all current models are readily available at https://github.com/AGhaderi/hDDM attention.

2.5. Multiple regression model

In this section, we explored whether both pairs of explanatory (independent) variables of {N2nc, Anc}, and {N2ni, Ani} could predict spatial prioritisation in the best-fitting model ($model_t$) described in the previous section, as well as behavioural performance (RT and accuracy). Multiple regressions using the Ordinary Least Squares (OLS) method were fit using a statistics module (statsmodels) in Python, with the following formula:.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon \tag{3}$$

Two regression model types were fit for each pair of regressors. In one model, X_1 was *N2nc* and X_2 was *Anc*, and in the other model *X*1 was *N2ni*, and X_2 was *Ani*. We allowed the dependent variable *y* to be either a direct estimate of behavioural performance (RT or accuracy) or an estimated non-decision time parameter from $model_t$, the best fitting DDM, per participant. In order to investigate the mechanism of spatial attention based on electrophysiological components in perceptual decision making, we applied the separate multiple regression models with prioritisation (informative spatial cueing) minus non-prioritisation condition (uninformative spatial cueing) dependent variables. That is, the dependent variables of these regressions were calculated by subtracting the non-decision time parameter (or mean RT or accuracy) of the informative spatial cueing condition from the non-decision time parameter (or mean RT or accuracy) of the uninformative spatial. The non-decision time parameters were estimated using mean posterior distribution for each participant from the best fitting model, $model_t$. The model regressors were the pairs of explanatory (independent) variables of *N2nc*, *Anc*, and *N2ni*, *Ani* previously mentioned. This resulted in six multiple regressions being run. Each multiple regression model was used to assess the relationship between electro-physiological correlates and latent processing of spatial prioritisation resulting from modelling parameters with *p*-values found for each beta coefficient.

We also calculated a Bayes factor (*BF*) for each model. *BF*s, which range from 0 to ∞, quantify the relative evidence for one model over another model (in our case, the null model with only an intercept/mean term). For example, if $BF_{10} = 5$, it means that the evidence is 5 times more probable under the alternative model than under a null model. When the BF equals 1, it indicates that both model models (i.e. the null and the alternative model) have equal probability. Moreover, a $BF_{10} > 1$ gives evidence in favour of the alternative model instead of the null model. If the $BF_{10} > 3$ or > 10, it shows moderate and strong evidence for the alternative model, respectively. We denote the BF with a subscript, BF_{10}, to clarify that this BF gives the probability in favour of the alternative model over the null model (Etz & Vandekerckhove, 2018). We calculated the BF_{10} in R using the BayesFactor package (Morey and Rouder).

2.6. Neurocognitive model fitting

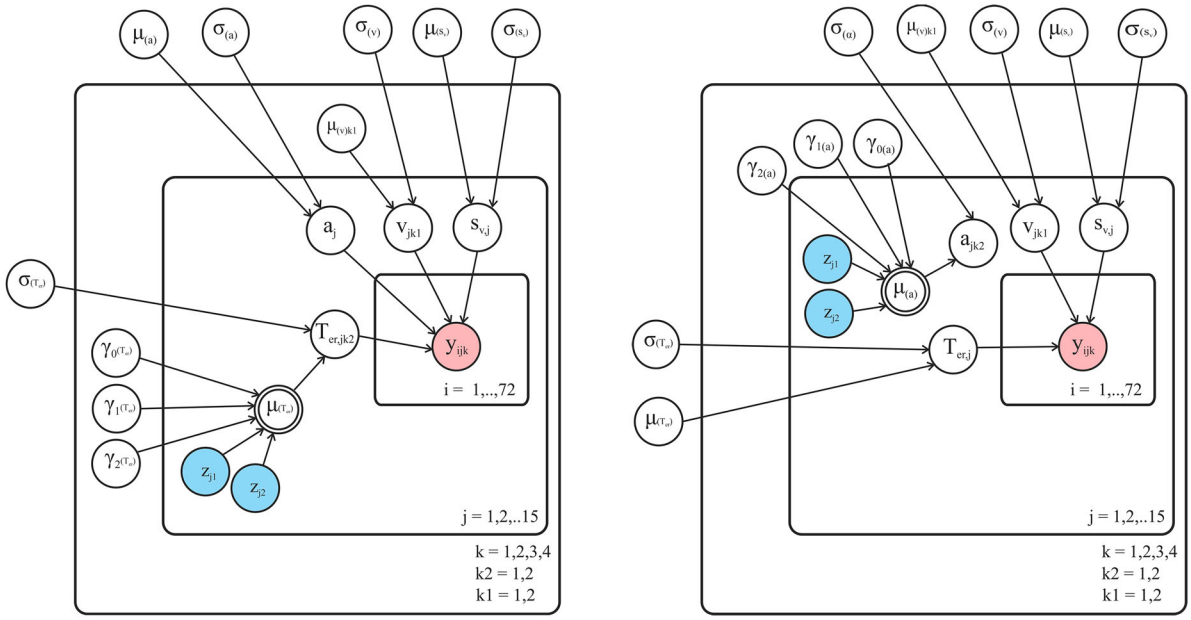
The analyses previously discussed are based on two-stage approaches to relate EEG measures and behavioral data. Specifically, the model parameters are estimated separately in hierarchical Bayesian methods and then parameter estimates (posterior means) are used as dependent variables in multiple linear regression models with EEG predictors. In order to ensure the robustness of the results to different methods, we also used *directed neurocognitive models* as a one-stage framework to constrain latent cognitive parameters to neural measures (Nunez et al., 2019; Palestro et al., 2018). Specifically, we created four directed neurocognitive models that included linear connections between EEG measures (central N2 subcomponent and alpha frequencies at the individual level) and the main parameters in a DDM.

For each directed neurocognitive model, we embedded neural covariates of N2nc and N2nc as well as Anc and Anci at the individual level into models. We then compared the contralateral and ipsilateral effects in the best model to show how EEG measures affect cognitive parameters. For each neurocognitive model, we assumed embedded linear regressions such that the difference of one parameter between prioritisation and non-prioritization can be driven by differences in EEG measures between the two conditions.

Neuro Model A (see Figure 3a) assumes that neural measures can predict the difference in non-decision time between two levels of spatial attention. The hyperpiors, priors, and likelihoods of the hierarchical Bayesian model are given as follows:.

$$v_{jk_1} \sim N(\mu_{(v)k_1}, \sigma^2_{(v)}), \quad \mu_{(v)k_1} \sim N(2, 3^2), \quad \sigma_{(v)} \sim \Gamma(1, 1), \tag{4a}$$

$$a_j \sim \mathcal{N}(\mu_{(a)}, \sigma^2_{(a)}), \quad \mu_{(a)} \sim \mathcal{N}(1, 2^2),$$



(a) Neuro Model A (b) Neuro Model C

Figure 3. Graphical representation of the hierarchical Bayesian models for based on the convention of Lee and Wagenmakers (Lee & Wagenmakers, 2014). (a) *Neuro Model A* assumes that prioritisation mainly manipulates non-decision time and neural measures predict the difference in non-decision time between two levels of spatial attention and (b) *Neuro Model C* assumes that prioritisation mainly manipulates boundary decision criteria and neural measures predict the difference in boundary decision between two levels of spatial attention. Nodes show manifest and random variables in the model and one-directional arrows indicate what variables affect another variable. In directed neurocognitive models, double-bounded nodes show deterministic parameters that do not need to estimate posterior. Index of i refers to experimental trials, indices of j refers to the number of participants, index of k indicates four conditions, and indices of $k1$ and $k2$ indicate coherence independent variable at two levels and spatial prioritisation independent variable at two levels. Manifest of x_{j1} and x_{j2} are contralateral and ipsilateral covariances at the individual level and y_{ijk} refers to response time multiply accuracy for each trail, each participant and each condition. [To view this figure in color, please see the online version of this journal.]

$$\sigma_{(a)} \sim \Gamma(1, 1), \tag{4b}$$

$$T_{er,j,2} - T_{er,j,1} \sim N(\gamma_{0(T_{er})} + \gamma_{1(T_{er})} \times x_{j,2} + \gamma_{2(T_{er})} \times x_{j,1}, \sigma^2_{(T_{er})}), \tag{4c}$$

$$\sigma_{(T_{er})} \sim \Gamma(.1, 1), \gamma_{0(T_{er})} \sim N(0, .2), \gamma_{1(T_{er})} \sim N(0, 1), \gamma_{2(T_{er})} \sim N(0, 1), \tag{4d}$$

$$s_{v,j} \sim N(\mu_{(s_v)}, \sigma^2_{(s_v)})\hat{I}(0,), \mu_{(s_v)} \sim N(1, 1^2), \sigma_{(s_v)} \sim \Gamma(1, 1), \tag{4e}$$

$$y_{ijk} \sim \text{Wiener}(T_{er,jk_2}, a, z, v_{jk_1}, s_{v,j}). \tag{4f}$$

where index i refers to experimental trials, indices j refer to participants, k refers to all conditions, k_1 refers to phase coherence and k_2 refers to the prioritisation condition. Parameters $T_{er,j,2}$ and $T_{er,j,1}$ indicate the non-decision time for non-prioritisation and prioritization conditions respectively. The parameter of z is fixed at 0.5 for all participants because we use accuracy-based implementation, v_i is the drift rate, $s_{v,j}$ is intrinsic trial-to-trial variability for drift rate, parameters of $x_{j,1}$ and $x_{j,2}$ indicate two neural measures for each participant. We run the above model two times: one where $x_{j,1} = N\ 2nc$ and $x_{j,2} = N\ 2ni$ and another where $x_{j,1} = Anc$ and $x_{j,2} = Ani$, $\gamma_{0(T_{er})}$ is intercept, $\gamma_{1(T_{er})}$ is the slope effect of the first neural measure as contralateral activity and $\gamma_{2(T_{er})}$is the slope effect of the second neural measure as ipsilateral activity. Note that the difference in posterior parameters between $\gamma_{1(T_{er})}$ and $\gamma_{2(T_{er})}$ can show the magnitude of each lateral activity related to spatial attention. To compute the uncertainty, we consider matched posterior samples of two effect parameters as a Bayesian probability approach. To do so, we compute the integral of the posterior distribution of $|\gamma_{1(T_{er})}| - |\gamma_{2(T_{er})}|$ from 0 to ∞.

Neuro Model B assumes that spatial attention manipulates evidence accumulation and neural measures also can predict the difference in evidence

accumulation between two levels of spatial attention such that the drift rate also changes with phase coherence. Therefore, we average the drift rate across coherence and then make the difference between spatial prioritisation as follows:.

$$(v_{j,1,1} + v_{j,2,1})/2 - (v_{j,1,2} + v_{j,2,2})/2 \sim \mathcal{N}(\gamma_{0(v)} + \gamma_{1(v)} \times x_{j,2} + \gamma_{2(v)} \times x_{j,1}, \sigma^2_{(v)}), \quad (5a)$$

$$\sigma_{(v)} \sim \Gamma(1, 1), \quad \gamma_{0(v)} \sim \mathcal{N}(2, 3), \gamma_{1(v)} \sim \mathcal{N}(0, 2), \gamma_{2(v)} \sim \mathcal{N}(0, 2), \quad (5b)$$

where the drift rate is considered as three dimensions for simplicity as (participants, coherence, prioritisation).

Neuro Model C (see Figure 3b) assumes that spatial attention manipulates the boundary decision and neural measures also can predict the difference in boundary decision between two levels of spatial attention as follows:.

$$a_{j,2} - a_{j,1} \sim N(\gamma_{0(a)} + \gamma_{1(a)} \times x_{j,2} + \gamma_{2(a)} \times x_{j,1}, \sigma^2_{(a)}), \sigma_{(a)} \sim \Gamma(1, 1), \gamma_{0(a)} \sim N(1, 2), \quad (6a)$$

$$\gamma_{1(a)} \sim \mathcal{N}(0, 2), \gamma_{2(a)} \sim N(0, 2). \quad (6b)$$

Neuro Model Full assumes that spatial attention manipulates all main parameters and neural measures also can predict the difference of them between the two levels of spatial attention. To fit each hierarchical model to both EEG and behavioural data simultaneously, we applied Stan as the probabilistic programming language and Markov Chain Monte Carlo (MCMC) sampler to estimate posterior parameters (Carpenter et al., 2017). For running each model, we used three MCMC chains and 6,000 samples for each chain, 2000 samples for the burn-in phase and a thinning parameter of 2 resulting in 6000 posterior samples for each parameter. Also, we used the Gelman-Rubin statistic, R-hat, to evaluate the convergence of posterior samples by comparing between-chain and within-chain variance. These models have R-hat statistics of less than 1.01 for all parameters. To assess model fits, we reported some model selection criteria to compare models and then select the best model: the lower the Widely Applicable Information Criterion (WAIC) is better, the higher the log pointwise predictive density (lppd) is better which indicates the log predictive accuracy of the model fits, and the effective number of parameters (PWAIC) indicates the complex model (Vehtari et al., 2017).

To see the robustness of the results of the neurocognitive models, we used a *non-parametric bootstrapping* method to re-sample behavioural and neural data. This approach is a flexible statistical technique to generate new data with replacement (Wagenmakers et al., 2004). We randomly took 30 participants as new data from *the original 14 participants* 30 times to improve model comparison inference and estimate better the gamma effect related to contralateral and ipsilateral activity.

3. Results

3.1. Behavioural result

Both coherence and spatial cueing were distinct experimental variables, which could have changed the quality of sensory information and the top-down attention to a specific visual field respectively. However, in order to understand the true effects of the experimental manipulation, the behavioural performance was analysed using three-way repeated measure ANOVA. For the ANOVA of mean response times for each participant, this ANOVA found no significant main effects of stimulus ($F(2, 13) = 0.78$, $p = 0.39$), but revealed significant main effects of spatial cueing **($F(2,13) = 11.83, p = 0.0044 < 0.01$)** and phase coherence ($F(2, 13) = 38.28$, $p < 0.0001$) (see Figure 1d). Specifically, the easier phase coherence condition and informative spatial cueing condition led to faster response times. Interaction between stimulus and coherence was also significant ($F(2, 13) = 8.54$, $p = 0.0119 < 0.05$), but the other interactions between variables were not statistically significant, i.e. interaction of spatial cueing and phase coherence ($F(2, 13) = 0.54$, $p = 0.47$) and interaction of spatial cueing and stimulus ($F(2, 13) = 0.13$, $p = 0.71$). Because of the significant interaction between coherence and stimulus, it was worthwhile to create DDMs describing response choices (e.g. face or car), and not just DDMs describing trial-by-trial accuracy (e.g. correct or incorrect), so that these effects can be observed with the drift rate. A three-way repeated measure ANOVA was also fit to mean accuracy for each participant. The only significant main effect ($F(2, 13) = 41.83$, $p < 0.0001$) (see Figure 1d) was that of coherence. The main effects of stimulus ($F(2, 13) = 0.84$, $p = 0.37$) and spatial cueing ($F(2, 13) = 2.12$, $p = 0.17$) were not significant. Similarly the interactions of spatial cueing and

phase coherence ($F(2,13) = 0.05$, $p = 0.81$), of spatial cueing and stimulus ($F(2, 13) = 2.9$, $p = 0.11$) and of coherence and stimulus ($F(2, 13) = 0.14$, $p = 0.71$) were not significant.

3.2. Cognitive modelling

To explore the underlying mechanisms of spatial attention in perceptual decision making, we fit the five hierarchical drift-diffusion models (HDDMs) previously mentioned. The convergence diagnostics of the best fitting model, $model_t$, are reported in the Supplementary Material of the article. MCMC sampling traces of the group posterior distributions showed converged chains during model fitting (see Figures 9–11). Furthermore, the convergence of the Markov chains was assessed by calculating the R-hat Gelman-Rubin statistic (Gelman & Rubin, 1992). It should be noted that $model_t$ showed superior convergence and the same stationary distribution with four chains, based on R-hat values under 1.0001 for the parameters (see Supplementary Material). Also, we compared posterior predictions from each model with the observed data. There was a near agreement between the observed data and the model predictions across conditions in all conditions, and the model is able to capture the overall shape of response times across the four different conditions (see Figure 12). Furthermore, the result of model comparison analysis reveals that $model_t$ provided the best model fit, in that it had the smallest DIC and largest R^2. The DIC should be noted that a difference in DIC scores of 15 and above is considered meaningful (Spiegelhalter et al., 2002; Spilcke-Liss et al., 2019). This suggests that the non-decision time parameter (T_{er}) is manipulated by spatial cueing, see Table 1, more so than other than other parameters.

We then utilised the results of $model_t$ to examine our hypotheses and identify the electrophysiological mechanisms of informative (contralateral and ipsilateral) and uninformative (neutral) prioritisation under decision-making. In order to assess parameters of the winning $model_t$ across conditions, we used a dependent Student's t-test on non-decision time and two-way repeated measure ANOVA on the drift-rate parameter. Mean non-decision processes of spatial cueing and non-spatial cueing were significantly different ($t = 3.34$, $p = 0.0053 < 0.01$), see Figure 4a. That is, we found that non-decision processes take less time when people receive an informative spatial cue than when they receive an ambiguous (uninformative) spatial cue. In addition, the main effect of coherence on drift rate was significant ($F(1, 13) = 75.11$, $p < 0.001$), while the stimulus had no effect on drift rate ($F(1, 13) = 0.3$, $p = 0.87$), see Figure 4b. There were also no significant interaction between stimulus and coherence on drift rate ($F(1, 13) = 0.64$, $p = 0.43$).

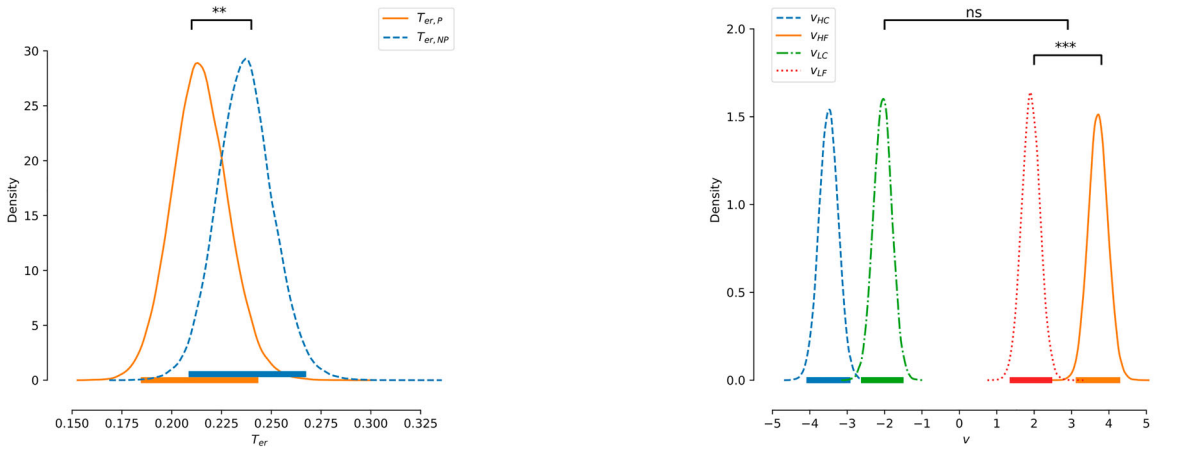
Correlation of the DDM and conventional outcome Parameters: To understand which characteristics of behavioural outcome are indicative of cognitive parameters, the correlations were found across participants between mean posterior values of non-decision time from $model_t$ and mean response times across trials. Matrix correlations are visualised in Figure 5. $RT_{er,P-Np}$ (response time of prioritisation minus response time of non-prioritisation) and $T_{er,P-NP}$ (non-decision time of prioritisation minus non-decision time of non-prioritisation) is highly positively correlated "0.47", indicating that larger $T_{er,P-NP}$ in the decision process is associated with longer responses RT_{P-NP}. Therefore, this indicates that RT changes relating to spatial attention are influenced by non-decision time.

3.3. Central N2 subcomponent and alpha lateralisation

The ERPs calculated from central electrodes C1/2/3/4 are shown in Figure 6. In this figure, *N2nc* and *N2ni* are highlighted as gray colours in Figure 6a and b respectively. These gray highlights are based on the significant t-test with FDR correction described

Table 1. Model comparison based on DIC for all model variants tested in this study and R^2 metrics for RT and accuracy in the group level. coher, coherence; stim, stimulus; spat, spatial attention; RT, response time; Acc, accuracy.

Model	$v \sim$ coher + stim	$v \sim$ spat	$Ter \sim$ spat	$z \sim$ spat	$a \sim$ spat	DIC	R^2(RT)	R^2(Acc)
$model\_p$	√					−3536.59	0.902	0.716
$model_v$	√	√				−3555.87	0.915	0.768
$model_t$	√		√			**−3610.76**	**0.940**	**0.888**
$model_z$	√			√		−3525.91	0.770	0.371
$model_a$	√				√	−3536.59	0.902	0.716



(a) Hierarchical non-decision time posterior distributions. (b) Hierarchical drift rate posterior distributions.

Figure 4. Posterior distribution of the group parameter non-decision time and drift rate from the $model_t$. (a), Non-decision was found to be different across the two types of cueing experimental conditions (t = 3.34, $p = 0.0053 < 0.01$). (b) The main effect of coherence on drift rate was significant ($F(1,13) = 75.11$, $p < 0.001$), but the main effect of stimulus on drift rate was not significant ($F(1,13) = 0.3$, $P = 0.87$). There was also no significant of interaction between stimulus and coherence of drift rate ($F(1,13) = 0.64$, $P = 0.43$). Each legend item refers to a different posterior distribution from $model_t$: $T_{er,P}$: non-decision time of prioritisation (informative spatial cueing), $T_{er,NP}$: non-decision time of non-prioritisation (uninformative spatial cueing), v_{HC}: drift rate of high coherence and car stimulus, v_{HF}: drift rate of high coherence and face stimulus, v_{LC}: drift rate of low coherence and car stimulus, v_{LF}: drift rate of low coherence and face stimulus. Asterisks indicate significant p-value as $^{**}p < 0.01$ and $^{***}p < 0.001$, and "np" indicates non-significant p-value. Thick horizontal lines are 95% of Highest Posterior Density (HPD). [To view this figure in color, please see the online version of this journal.]

previously. Figure 6a shows the grand-average across participants of contralateral and neutral ERPs and the difference between them, time-locked to cue onset. FDR correction over samples of the time window found two distinct N2nc sub-components at central electrode sites with (the mean $p = 0.018 < 0.05$ for 190–235 msec and the mean $p = 0.024 < 0.05$ for time windows 260–315 msec after stimulus onset). Figure 6b presents the grand-average of signal ipsilateral and neutral and also ipsilateral minus neutral. As a result, the FDR correction over samples of the

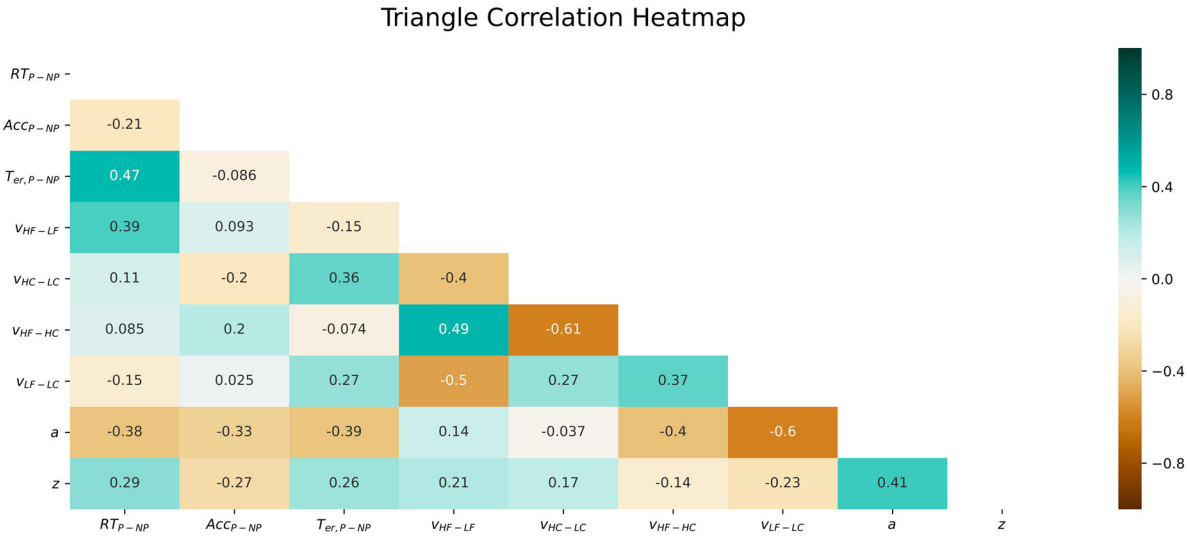
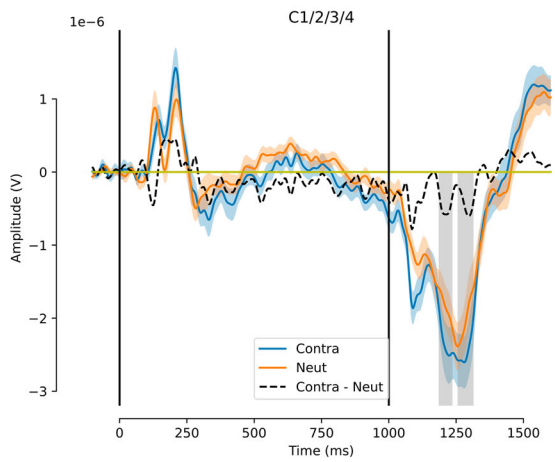
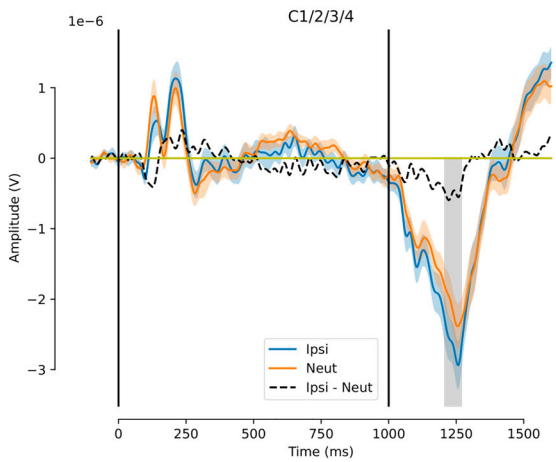


Figure 5. The heatmap matrix correlation of the participant-level parameters across conditions. $RT_{P\ -NP}$: RT of prioritisation minus RT of Non-prioritisation, $Acc_{P\ -NP}$: Accuracy of prioritisation minus Accuracy of Non-prioritisation, $T_{er,P-Np}$: $T_{er,P}$ minus $T_{er,NP}$, V_{HF-LF}: V_{HF} minus V_{LF}, V_{HC-LC}: V_{HC} minus V_{LC}, V_{HF-HC}: V_{HF} minus V_{HC}, V_{LF-LC}: V_{LF} minus V_{LC}. [To view this figure in color, please see the online version of this journal.]



(a) Contralateral versus neutral ERP at central electrodes.



(b) Ipsilateral versus neutral ERP at central electrodes.

Figure 6. N2nc and N2ni subcomponents at central electrodes C1/2/3/4 highlighted by the shaded gray bars. (a) The grand-average across participants of contralateral and neutral ERPs as well as the N2nc (contralateral minus neutral) are displayed. The significant windows in gray are based on t-tests with $\alpha = 0.05$ and an FDR correction within a time window from 150 msec to 500 msec after stimulus onset is illustrated by the gray colour. Specifically, the gray colour indicates the mean $p < 0.05$ for 190–235 msec and the mean $p < 0.05$ for 260–315 msec windows after stimulus onset (i.e. 1190–1235 msec 1260–1315 after cue onset). (b) The grand-average ipsilateral and neutral ERPs as well as the N2ni (ipsilateral minus neutral) are displayed. In addition, the significant windows with FDR correction are illustrated by the gray colour. Specifically the gray colour indicates the mean $p < 0.05$ for 210–270 msec time window after stimulus onset (i.e. 1210–1270 after cue onset). In each figure, the first and second vertical lines show the time point of the spatial cue (single-sided or double-sided arrows) and the stimulus presence (face or car) respectively. The blue and orange shaded regions identify the standard error of each ERP. [To view this figure in color, please see the online version of this journal.]

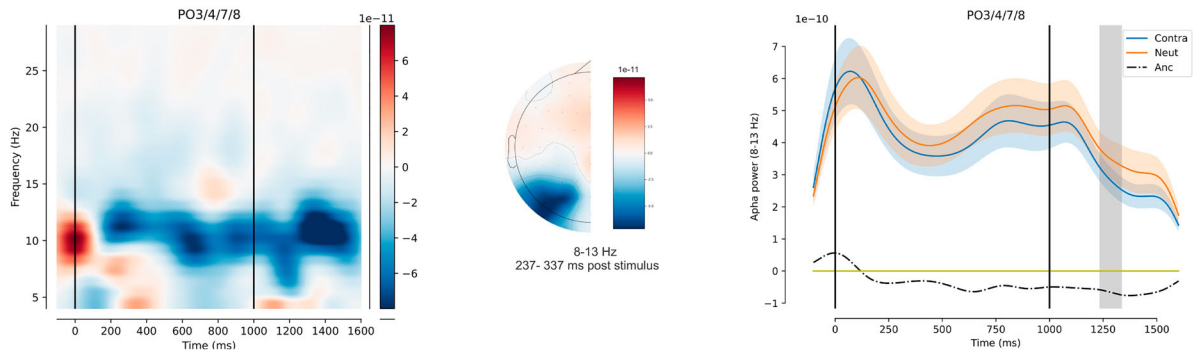
time window revealed an N2ni subcomponent at the central electrode sites with (mean $p = 0.022 < 0.05$ for time window 210–270 msec after stimulus onset).

Figure 7 presents a grand-average of the single-trial time–frequency representation asymmetric of alpha power from 8 Hz to 13 Hz at posterior-occipital sites (PO3/4/7/8). At first, we considered the ± 50 msec around 50% FAL (237-337 msec after stimulus onset) between −100 msec to 600 msec after stimulus in the Anc signal. Then, by applying a one-sample t-test in the mean portion of Anc, we found a significant effect with (t = 2.59, $p = 0.022$ ¡ 0.05), as shown in Figure 7a with the gray colour. Also, we examined the same analysis in the Ani signal and a time window (155-255 msec after stimulus onset) was found to be significant. Therefore, using a one-sample t-test in the mean portion of Ani, we had a significant effect with (t = 2.27, $p = 0.040$ ¡ 0.05) which is shown by the gray colour, see Figure 7b.

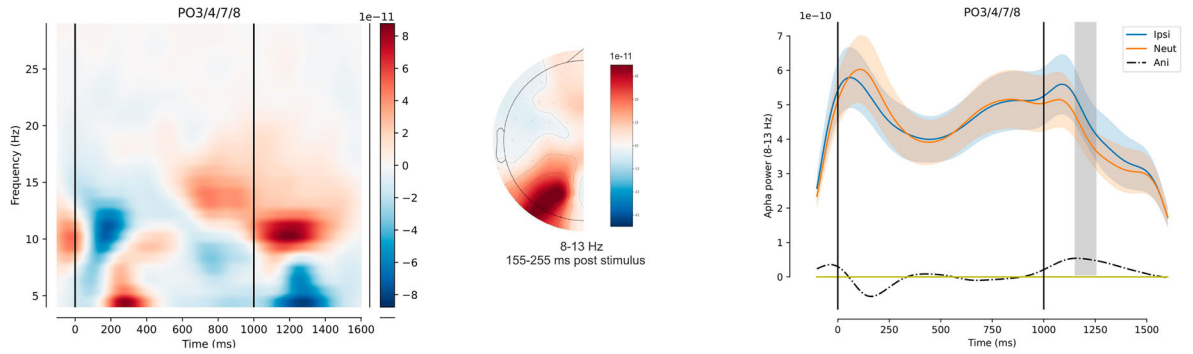
3.4. Regression analysis

Tables 2 and 3 show the result of model fitting for both of contralateral versus neutral component and ipsilateral versus neutral component as the difference between informative versus uninformative neutral activity. Mean response times were predicted by both the *N2nc* with time window 260–315 msec post-stimulus and *Anc* with the time window 237–337 msec post-stimulus ($R^2 = 0.679$, $F(2, 13) = 11.66$, $p = 0.00192 < 0.01$). There was no significant effect for *N2nc* with a time window 190–235 post-stimulus, so we disregarded it from further analysis. In this regression model, the main effect of *Anc* was highly significant ($t = 3.377$, $p = 0.006 < 0.01$) and the main effect of N2nc was not significant ($t = 1.549$, $p = 0.150$). However, the Pearson correlation coefficient disclosed the relationship between mean RT and N2nc ($r = 0.589$, $p = 0.026 < 0.05$).

Non-decision time differences explained by *N2nc* and *Anc* ($R2 = 0.428$, $F(2, 13) = 4.114$, $p = 0.0464 < 0.05$). In addition, the main effect of N2nc was significant (t = 2.746, $p = 0.019$ ¡0.05), and the main effect of Anc was not significant ($t = -0.504$, $p = 0.624$). However, accuracy was not predicted by N2nc and Anc ($R^2 = 0.123$, $F(2, 13) = 0.7701$, $p = 0.486$).



(a) Contralateral versus neutral power at posterior-occipital electrodes.



(b) Ipsilateral versus neutral power at posterior-occipital electrodes.

Figure 7. Time–frequency lateralisation. (a) The grand-average of contralateral minus neutral single-trial time–frequency wavelets is indicated at electrodes PO3/4/7/8 on the left-hand side. In addition, the mean of contralateral and neutral alpha power from 8 Hz to 13 Hz as well as the outcome of (contralateral minus neutral) are displayed on the right-hand. The ± 50 msec around of 50% FAL on the time window from −100 msec to 600 msec of the stimulus onset was achieved in the 237–337 msec window after the stimulus onset. The topographic map is based on differences in contralateral minus neutral alpha oscillation in the significant time window. (b) The grand-average of ipsilateral minus neutral single-trial time–frequency wavelets are also shown at the posterior-occipital electrodes on the left-hand side. Moreover, the mean of contralateral and neutral alpha power, as well as the outcome of (ipsilateral minus neutral), are displayed on the right-hand side. Therefore, the result of FAL was achieved in 155–255 msec after the stimulus onset. The topographic map is based on differences of ipsilateral minus neutral alpha oscillation in the significant time window. The first and second vertical lines show in the right and left plots indicate cue and stimulus occurrences respectively, and shaded regions identify the standard error. [To view this figure in color, please see the online version of this journal.]

For regressions with ipsilateral versus neutral regressors, there were also no significant models, see Table 3.

3.5. Neurocognitive models

We seek to discover which DDM parameters as well as which contralateral and ipsilateral activity was mostly manipulated by prioritisation during face-car perceptual decision making. Therefore, we used model selection criteria and the result of four neurocognitive models to find out the best model fit and the difference of magnitude effect parameters. Based on Table 4, *Neuro Model A* was selected as the best model since it has lower WAIC, upper LPPD and upper ELPD. The result of model comparison over neurocognitive models matches the result of cognitive models in previous sections in which non-decision time is pivotal in manipulating spatial attention. However, neurocognitive models provide some more information about the relationship between neural covariance and cognitive parameters related to spatial prioritisation. Figure 8 shows parameter posteriors of hierarchical Bayesian *Neuro Model A* between two main independent variables which is the 95% Bayesian credible interval (BCI). Bayesian inference shows that there is a 72%

Table 2. Separate multiple linear regressions between the dependent variables (non-decision time of prioritisation minus non-prioritisation, RT and accuracy) and predictors (N2nc and Anc), FDR correction for two different N2nc, t-test statistics for each Beta parameters, confidence intervals, and standard errors.

Output	$T_{er,P-NP}$		RT_{P-NP}		Acc_{P-NP}	
Predictors	β (SE) [95% CI]	t	β (SE) [95% CI]	t	β (SE) [95% CI]	t
Intercept	−0.0046(0.009) [−0.024 0.015]	−0.523 $p = 0.612$	−3.5859(6.836) [−18.63 11.46]	t = −0.525 $p = 0.610$	0.1295(1.249) [−2.621 2.880]	t = 0.104 $p = 0.919$
N2nc	4.712e + 04*(1.72e + 04) [9354.840 8.49e + 04] $BF_{10} = 4.361$	t = 2.746 $p = 0.019$	2.048e + 07(1.32e + 07) [8.62e + 06 4.96e + 07] $BF_{10} = 2.674$	t = 1.549 $p = 0.150$	−2.999e + 06(2.42e + 06) [−8.32e + 06 2.32e + 06] $BF_{10} = 0.684$	t = −1.241 $p = 0.241$
Anc	−3.485e + 07(6.92e + 07) [−1.87e + 08 1.17e + 08] $BF_{10} = 0.516$	t = −0.504 $p = 0.624$	1.8e + 11**(5.33e + 10) [6.27e + 10 2.97e + 11] $BF_{10} = 29.217$	t = 3.377 $p = 0.006$	5.19e + 09(9.74e + 09) [−1.63e + 10 2.66e + 10] $BF_{10} = 0.446$	t = 0.533 $p = 0.605$
R^2	0.428		0.679		0.123	
F-statistic	F(2,13) = 4.114, $p = 0.0464$ $BF_{10} = 2.036$		F(2,13) = 11.66, $p = 0.00192$ $BF_{10} = 21.83$		F(2,13) = 0.7701, $p = 0.486$ $BF_{10} = 0.415$	

$T_{P-NP} = T_{er}(yes) - T_{er}(no)$, $RT\_P - N\_P = RT(P) - RT(NP)$, $Acc\_\{P - NP\}$ = accuracy(P) - accuracy(NP), SE = standard error, CI = confidence interval. R^2 and regression coefficients are given for separate linear regression models, yes = prioritisation, no = no prioritisation. Asterisks indicate significant p-value as $^*p < 0.05$, $^{**}p < 0.01$.

Table 3. Separate multiple linear regressions between the dependent variables (non-decision time of prioritisation minus non-prioritisation, RT and accuracy) and predictors (N2ni and Ani), FDR correction for two different N2nc, t-test statistics for each Beta parameters, confidence intervals, and standard errors.

Output	$T_{er,P-NP}$		RT_{P-NP}		Acc_{p-NP}	
Predictors	β (SE) [95% CI]	t	β (SE) [95% CI]	t	β (SE) [95% CI]	t
Intercept	−0.0194(0.013) [−0.047 0.008]	t = −1.548 $p = 0.150$	−11.6572(10.140) [−33.976 10.662]	t = 1.1700(1.326) $p = 0.275$	t = 0.882 [−1.749 4.089]	$p = 0.397$
N2ni	8286.7188(2.52e + 04) [−4.71e + 04 6.36e + 04] $BF_{10} = 0.466$	t = 0.329 $p = 0.748$	8.253e + 06(2.04e + 07) [−3.66e + 07 5.32e + 07] $BF_{10} = 1.333$	t = 0.405 $p = 0.694$	1.82e + 06(2.67e + 06) [−4.05e + 06 7.69e + 06] $BF_{10} = 0.456$	t = 0.682 $p = 0.509$
Ani	1.225e + 07(1.22e + 08) [−2.56e + 08 2.81e + 08] $BF_{10} = 0.450$	t = 0.100 $p = 0.922$	−1.643e + 11(9.89e + 10) [−3.82e + 11 5.35e + 10] $BF_{10} = 3.312$	t = −1.661 $p = 0.125$	1.703e + 10(1.29e + 10) [−1.14e + 10 4.55e + 10] $BF_{10} = 0.698$	t = 1.316 $p = 0.215$
R^2	0.012		0.387		0.141	
F-statistic	F(2,13) = 0.0674, $p = 0.94$ $BF_{10} = 0.272$		F(2,13) = 3.469, $p = 0.0679$ $BF_{10} = 1.560$		F(2,13) = 0.7701, $p = 0.434$ $BF_{10} = 0.447$	

$T\_\{P - NP\} = T\_\{er\}(P) - Ter(NP)$, $RT\_\{P - NP\} = RT(P) - RT(NP)$, $Acc\_\{P - NP\}$ = accuracy(P) - accuracy(NP), SE = standard error, CI = confidence interval. R^2 and regression coefficients are given for separate linear regression models, yes = prioritisation, no = no prioritisation. Asterisks indicate significant p-value as $^*p < 0.05$, $^{**}p < 0.01$.

(see Figure 8d probability of the difference in slope effects greater than 0 related to contralateral and ipsilateral ERP for time measurement 260–315 msec and 62% or time measurement 190–235 msec after stimulus appearance, and also 67% probability of the difference related to contralateral and ipsilateral alpha frequencies. These results identify that contralateral activity is more influential than ipsilateral activity in manipulating spatial attention. For each model, R-hat was less than 1.01 for all parameters.

Table 4. Comparing hierarchical Bayesian models based on Model Comparison criteria for contralateral and ipsilateral ERP and alpha. Bold text indicates the best model fit statistic between four models.

Measure	Criterion	Neuro Model A	Neuro Model B	Neuro Model C	Neuro Model Full
Amplitude	WAIC	**−2840**	−2781	−2793	−2826
	LPPD	**1516**	1497	1485	1546
	ELPD	**1413**	1388	1394	1410
	PWAIC	95	107	89	133
Frequency	WAIC	**−2843**	−2780	−2794	−2827
	LPPD	**1516**	1496	1485	1547
	ELPD	**1418**	1389	1395	1409
	PWAIC	94	106	89	133

The result of using non-parametric bootstrap shows that the above results are robust across some boot-straps. In fact, the probability of the mean difference slope effects greater than zero for time measurement 260–315 msec across 20 boot-straps is 80%, for time measurement 190–235 msec

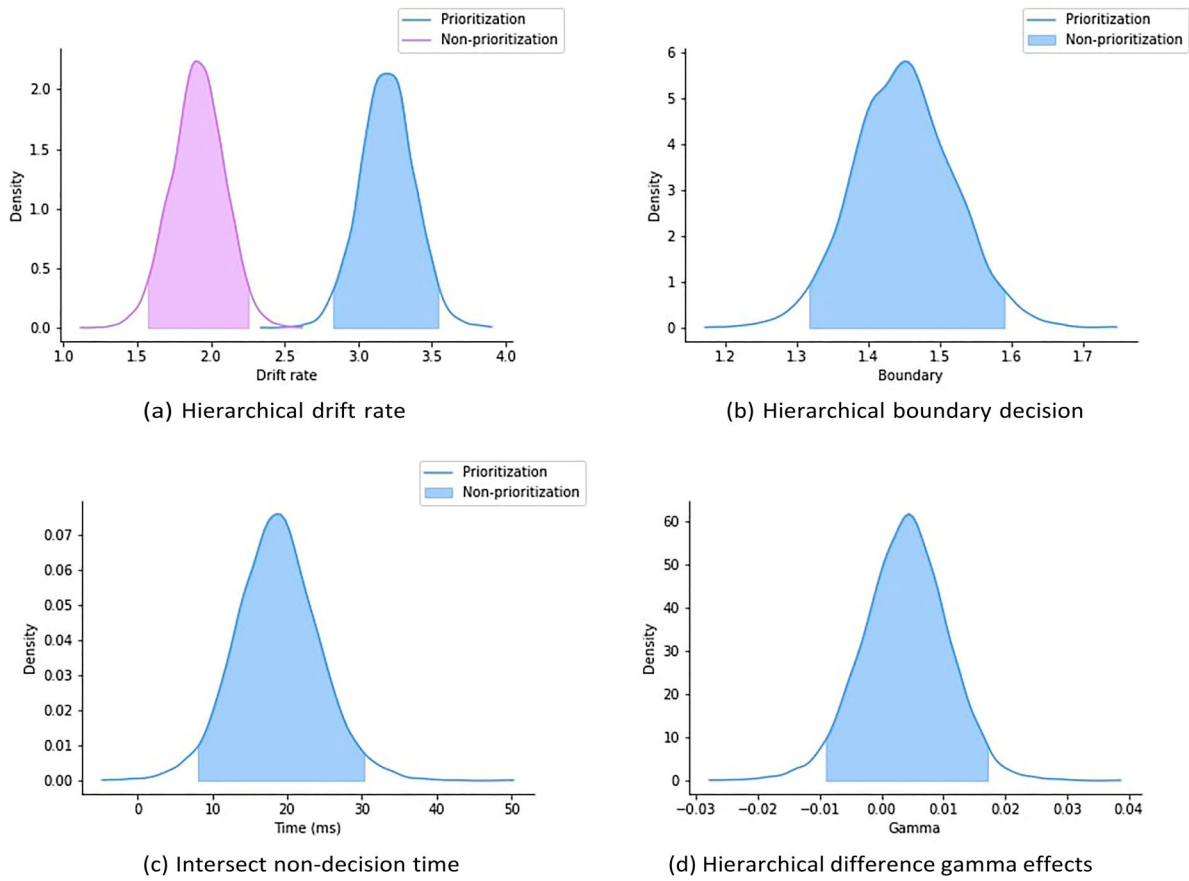


Figure 8. Parameter posteriors of *Neuro Model A* for contralateral and ipsilateral ERP. The shaded area represents the 95% Bayesian credible interval (BCI) of posteriors. [To view this figure in color, please see the online version of this journal.]

is 69%, and also the probability of the difference of alpha frequencies is 71%. For each bootstrap, R-hat was less than 1.01 for all parameters.

4. Discussion

This work investigated how spatial prioritisation cues lead to faster response times than non-spatial prioritisation cues in perceptual decision making tasks. In order to explain the elaborate framework of the observed behaviour, cognitive process, and observed electrophysiology during the top-down intention, we focused on the neural contribution of post-stimulus N2 subcomponents and alpha power. Contralateral minus uninformative N2 subcomponents and alpha power predicted the behavioural data and relevant DDM parameters. We differentiated between five competing cognitive models: 1. Spatial attention may not shift resources related to decision-making. 2. Spatial attention might affect the information accumulation process. 3. Spatial attention could be related to the speed of decoding or motor response (non-decision time). 4. Spatial attention could lead to a bias toward the presented stimulus. 5. Spatial attention could shift the amount of information that is needed to decide. We found Model 3 ($model_t$) best described the data as measured by DIC and R^2. In addition, we tested whether lateralised central sites relating to motor response and parieto-occipital relating to decoding stimulus reflected individual differences in model parameters. Specifically, we investigated the contribution of alpha power lateralisation and N2 sub-component amplitudes to individual differences in spatial attention in perceptual decision making. Our study revealed that both contralateral N2 and alpha power regressors could predict the non-decision parameter of the $model_t$ and RT associated with the spatial cue, see Table 2. Whereas, ipsilateral measures could not predict non-decision parameter of the $model_t$ and RT associated with the spatial cue, see Table 3.

We found that only the difference between contralateral and neutral components which could

predict the difference in non-decision times ($T_{er,P-NP} = T_{er}(P) - T_{er}(NP)$) and RT ($RT_{P-NP} = RT(P) - RT(NP)$) between prioritisation and non-prioritisation. This finding confirms that the mechanism of contralateral oscillations to the attended spatial location is more important than ipsilateral oscillation. Most previous works (Cespon et al., 2016; Gamble & Luck, 2011; Klatt et al., 2020; Luck, 2014) have concentrated on the difference between contralateral and ipsilateral deflections and have not explored which signals, by themselves, could predict relevant the DDM parameters associated with spatial attention. In the current study, we also found a significant relationship between non-decision time and contralateral minus neutral activity (Klatt et al., 2020; Praamstra, 2006).

Also, we tested four neurocognitive models to relate neural covariates and cognitive latent variable parameters simultaneously in model fits. We found that a model which varies with the non-decision time parameter is the best model based on PWAIC, LPPD and ELPS criteria. This result matches the result of cognitive models. Moreover, we used contralateral and ipsilateral N2 and alpha embedded in regression form in neurocognitive models and then saw that the magnitude gamma effect of contralateral activity is larger than the gamma effect of ipsilateral activity for both N2 and alpha band frequencies. Therefore, both multiple regression and neurocognitive models emphasise the role of contralateral activity in spatial attention as well as both cognitive and neurocognitive models emphasise the role of non-decision time to manipulate spatial attention.

4.1. N2nc and Anc band predict response time

Behavioural data revealed a significant effect of spatial prioritisation relating to non-prioritisation. Therefore, in order to decipher the mechanism of the difference, a multiple regression model using N2nc (the time measurement 260–315 msec) and Anc band as regressors was fit. This model could explain the difference RT between prioritisation and non-prioritisation, see Table 2. The effect of N2nc and Anc coefficients were displayed. Also, the components of N2ni and Ani together could not explain the variance of the difference RT, see Table 3. This result shows that contralateral activity could encode the different spatial cues in comparison with ipsilateral activity. According to post-stimulus of Figure 6, both N2 subcomponent contralateral and ipsilateral ERP were more negative than neutral ERP in line with latter findings (Gamble & Luck, 2011; Praamstra, 2006).

Post-stimulus contralateral alpha power was more negative than neutral alpha power, but ipsilateral alpha power was more positive than neutral alpha power, see Figure 7. Negative Anc suggests that the alpha power contralateral to the attend visual field decreased because participants' brains were made more sensitive to decoding the stimulus from the attended area. On the other hand, the positive Ani indicates that the alpha power is increased on the other side of the present field of view because the participants' brains are thought to seek to attenuate the non-attended spatial location. (Ikkai et al., 2016; Rihs et al., 2007). In our research, we separated these two mechanisms in order to understand the role of spatial attention encoded by EEG on perceptual decision making. The regression and statistical analyses proved that the role Anc and N2nc are more significant to spatial attention. That is N2nc and Anc led to more differences in RT associated with spatial attention.

4.2. N2nc amplitude predicts non-decision times

We also explored the role of alpha lateralisation and N2nc in visual perceptual decision making assuming a cognitive model, a DDM. We found that the difference in RT of prioritisation (informative cue conditions) and non-prioritisation (uninformative cue conditions) originated from the difference in non-decision time parameter based on the nested model comparisons. We show that N2nc and Anc predicted the non-decision parameter based on the nested model comparison. However, the main effect of N2nc was significant, but Anc was not significant. N2ni and Ani could not predict non-decision time parameter, see Table 3 Thus the N2nc might reflect the non-decision parameter that influences RT, while the Anc might reflect RTs without a clear cognitive parameter correlate.

The role of N2nc was easier to interpret as mediating the non-decision parameter difference. It also shows that the contralateral activity was superior to elucidating the effect of the spatial cue in comparison with ipsilateral activity. Therefore, the larger differences in non-decision time measured by

N2nc, the larger the differences in non-decision time relating to spatial attention.

5. Conclusion and future works

The non-decision time parameter reflects the spatial cue in perceptual decision making based on our findings. This result is supported by both cognitive and neurocognitive models such that a model with varying non- decision time to spatial prioritisation is the best model based on model selection criteria. We also found that the amplitude of N2nc at the central site predicted RT as well as the non-decision time due to spatial prioritisation across participants. The Anc power oscillations at parieto-occipital electrodes only predicted RT of spatial prioritisation across participants and did not predict non-decision time across participants. In addition, neither the N2ni amplitudes nor Ani power could predict RT and non-decision time differences between spatial prioritisation and non-prioritisation across participants. Therefore, the role of contralateral activity is shown to allocate the brain's resources in order to aid the decision making process, in comparison to ipsilateral activity. Larger N2nc led to larger differences in non-decision time associated with spatial attention, while both larger N2nc and Anc differences led to larger differences in mean RTs associated with spatial attention. In addition, the non-decision time parameter might explain the modulation of the attentional orienting during decision making process. We also found that neurocognitive models can be a good alternative to two-stage inference where a regression model is used to bridge the gap between cognitive parameters and neural messages. The result of neurocognitive models shows that contralateral activity containing N2 and alpha oscillations is more pivotal than ipsilateral activity in spatial prioritisation. This kind of analysis used in this work can be developed and used by other researchers to track the underlying latent process in spatial attention in different paradigms.

For future work, different joint-modelling techniques such as the Integrative, Directed, and Covariance approaches should be used in order to provide moment-by-moment descriptions of the neural data and a trial-by-trial descriptions of behavioural data. Moreover, combining latent processes and the dynamics underlying varied modalities (e.g. behavioural data, EEG, fMRI and so forth) could yield more insight into the neurocognitive role of spatial attention (Bahg et al., 2020; Nunez et al., 2019; Palestro et al., 2018; Turner et al., 2017, 2019b, 2016).

Acknowledgement

The authors are very grateful to Dirk Ostwald and his colleagues that made the data available.

Consent to participate

As noted in (Georgie et al., 2018), all participants gave informed consent for participation in the study.

Consent for publication

All authors consent to the publication of this manuscript.

Contributions

JAR developed the study concept. All authors contributed to the study design. AGK implemented and conducted the simulation experiments. All the authors contributed to analysing and interpreting the results. AGK drafted the manuscript. KP, RE, MDN and JAR provided critical revisions. All the authors approved the final version of the manuscript for submission.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

MDN was supported in part by US NSF grants (1658303 and 1850849). Also, JAR was supported in part by IRAN Cognitive Sciences and Technologies Council & Technologies Council (grant number 9311), and Iran National Science Foundation, INSF (grant number 99010447 and 4003215).Center for Hierarchical Manufacturing, National Science Foundation.

Ethics approval

As noted in (Georgie et al., 2018), the experiments were approved by the University of Birmingham Ethical Review Committee.

ORCID

Amin Ghaderi-Kangavari http://orcid.org/0000-0002-5666-5664
Kourosh Parand http://orcid.org/0000-0001-5946-0771

Reza Ebrahimpour http://orcid.org/0000-0002-7013-8078
Michael D. Nunez http://orcid.org/0000-0002-9965-6282
Jamal Amani Rad http://orcid.org/0000-0002-7322-7412

Data availability statement

All data used in this research is publicly available in the open science framework (https://osf.io/q4t8k/). Also, the hDDM implementation code and simulation scripts for all current models are readily available at https://github.com/AGhaderi/hDDM attention.

References

Amenedo, E., Lorenzo-Lopez, L., & Pazo-alvarez, P. (2012). Response processing during visual search in normal aging: The need for more time to prevent cross talk between spatial attention and manual response selection. *Biological Psychology*, *91*(2), 201–211. https://doi.org/10.1016/j.biopsycho.2012.06.004

Bahg, G., Evans, D. G., Galdo, M., & Turner, B. M. (2020). Gaussian process linking functions for mind, brain, and behavior. *Proceedings of the National Academy of Sciences*, 29398–29406. https://doi.org/10.1073/pnas.1912342117

Bernier, P. M., Whittingstall, K., & Grafton, S. T. (2017). Differential recruitment of parietal cortex during spatial and non-spatial reach planning. *Frontiers in Human Neuroscience*, *11*, 249. https://doi.org/10.3389/fnhum.2017.00249

Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, *112*(1), 117. https://doi.org/10.1037/0033-295X.112.1.117

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). *Stan*: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Cespon, J., Galdo-Alvarez, S., & Diaz, F. (2016). Cognitive control activity is modulated by the magnitude of interference and pre-activation of monitoring mechanisms. *Scientific Reports*, *6*(1), 1–11. https://doi.org/10.1038/srep39595

Choi, W., & Paik, S. B. (2019). Occurrence of the potent mutagens 2-nitrobenzanthrone and 3-nitrobenzanthrone in fine airborne particles. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-018-37186-2

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *The Journal of Neuroscience*, *32*(11), 3612–3628. https://doi.org/10.1523/JNEUROSCI.4010-11.2012

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin & Review*, *25*(1), 5–34. https://doi.org/10.3758/s13423-017-1262-3

Evans, N. J., Bennett, A. J., & Brown, S. D. (2019). Optimal or not; depends on the task. *Psychonomic Bulletin & Review*, *26*(3), 1027–1034. https://doi.org/10.3758/s13423-018-1536-4

Evans, N. J., & Brown, S. D. (2017). People adopt optimal policies in simple decision-making, after practice and guidance. *Psychonomic Bulletin & Review*, *24*(2), 597–606. https://doi.org/10.3758/s13423-016-1135-1

Evans, N. J., Hawkins, G. E., & Brown, S. D. (2020). The role of passing time in decision-making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 316. https://doi.org/10.1037/xlm0000725

Evans, N. J., & Wagenmakers, E. J. (2020). *Evidence accumulation models: Current limitations and future directions*. The Quantitative Methods for Psychology.

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, *26*(4), 1099–1121. https://doi.org/10.3758/s13423-018-1554-2

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, *67*(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2017). Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention. *Psychological Science*, *28*(7), 929–941. https://doi.org/10.1177/0956797617699167

Frank, M. J. (2015). Linking across levels of computation in model-based cognitive neuroscience. In *Forstmann, B., Wagenmakers, EJ. (eds) An Introduction to Model-Based Cognitive Neuroscience* (pp. 159–177). Springer New York. https://doi.org/10.1007/978-1-4939-2236-9_8

Gamble, M. L., & Luck, S. J. (2011). N2ac: An ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology*, *48*(8), 1057–1068. https://doi.org/10.1111/j.1469-8986.2010.01172.x

Gamerman, D., & Lopes, H. F. (2006). *Markov chain monte carlo: Stochastic simulation for Bayesian inference*. Taylor and Francis.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Georgie, Y. K., Porcaro, C., Bagshaw, A. P., Mayhew, S. D., & Ostwald, D. (2018). A perceptual decision making eeg / fmri data set. bioRxiv.

Gherman, S., & Philiastides, M. G. (2018). Human VMPFC encodes early signatures of confidence in perceptual decisions. *eLife*, *7*, 1–18. https://doi.org/10.7554/eLife.38293

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, *4*(6), 634–645. https://doi.org/10.1038/s41562-020-0822-0

Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision

making. *eLife*, *7*, e39659. https://doi.org/10.7554/eLife.39659

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389. https://doi.org/10.1037/0096-3445.136.3.389

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & H¨am¨al¨ainen, M. (2013). MEG and EEG data analysis with mne-python. *Frontiers in Neuroscience*, *7*, 1–13. https://doi.org/10.3389/fnins.2013.00267

Haegens, S., Luther, L., & Jensen, O. (2012). Somatosensory anticipatory alpha activity increases to suppress distracting input. *Journal of Cognitive Neuroscience*, *24*(3), 677–685. https://doi.org/10.1162/jocn_a_00164

Ikkai, A., Dandekar, S., & Curtis, C. E. (2016). Lateralization in alpha-band oscillations predicts the locus and spatial distribution of attention. *PLoS One*, *11*(5), e0154796. https://doi.org/10.1371/journal.pone.0154796

Imani, E., Harati, A., Pourreza, H., & Goudarzi, M. M. (2021). Brain-behavior relationships in the perceptual decision-making process through cognitive processing stages. *Neuropsychologia*, *155*, 107821. https://doi.org/10.1016/j.neuropsychologia.2021.107821

Klatt, L. I., Schneider, D., Schubert, A. L., Hanenberg, C., Lewald, J., Wascher, E., & Getzmann, S. (2020). Unraveling the relation between EEG correlates of attentional orienting and sound localization performance: A diffusion model approach. *Journal of Cognitive Neuroscience*, *32*(5), 945–962. https://doi.org/10.1162/jocn_a_01525

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, *3*, 193. https://doi.org/10.3389/fpsyg.2012.00193

Kvam, P. D. (2019). A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *Journal of Mathematical Psychology*, *91*, 14–37. https://doi.org/10.1016/j.jmp.2019.03.001

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *Journal of Mathematical Psychology*, *55*(1), 1–7. https://doi.org/10.1016/j.jmp.2010.08.013

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Li, Y., Lou, B., Gao, X., & Sajda, P. (2013). Post-stimulus endogenous and exogenous oscillations are differentially modulated by task difficulty. *Frontiers in Human Neuroscience*, *7*, 1–10. https://doi.org/10.3389/fnhum.2013.00009

Loughnane, G. M., Newman, D. P., Bellgrove, M. A., Lalor, E. C., Kelly, S. P., & O'Connell, R. G. (2016). Target selection signals influence perceptual decisions by modulating the onset and rate of evidence accumulation. *Current Biology*, *26*(4), 496–502. https://doi.org/10.1016/j.cub.2015.12.049

Luck, S. J. (2014). An introduction to the event-related potential technique. *MIT press*.

Mallahi-Karai, K., & Diederich, A. (2019). Decision with multiple alternatives: Geometric models in higher dimensions — the cube model. *Journal of Mathematical Psychology*, *93*, 102294. https://doi.org/10.1016/j.jmp.2019.102294

Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D., & Wagenmakers, E. J. (2013). Bayesian parametric estimation of stop-signal reaction time distributions. *Journal of Experimental Psychology: General*, *142*(4), 1047. https://doi.org/10.1037/a0030543

Morey, R. D., & Rouder, J. N. Bayesfactor: Computation of bayes factors for common designs (r package version 0.9.12-4.2). https://cran.r-project.org/package=bayesfactor.

Nunez, M. D., Gosai, A., Vandekerckhove, J., & Srinivasan, R. (2019). The latency of a visual evoked potential tracks the onset of decision making. *Neuroimage*, *197*, 93–108. https://doi.org/10.1016/j.neuroimage.2019.04.052

Nunez, M. D., Nunez, P. L., & Srinivasan, R. (2016). Electroencephalography (EEG): Neurophysics, experimental methods, and signal processing. In H. Ombao, M. Linquist, W. Thompson, & J. Aston (Eds.), *Handbook of neuroimaging data analysis* (pp. 175–197). Chapman & Hall/CRC.

Nunez, M. D., Srinivasan, R., & Vandekerckhove, J. (2015). Individual differences in attention influence perceptual decision making. *Frontiers in Psychology*, *6*, 18. https://doi.org/10.3389/fpsyg.2015.00018

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. *Journal of Mathematical Psychology*, *76*, 117–130. https://doi.org/10.1016/j.jmp.2016.03.003

Olianezhad, F., Zabbah, S., Tohidi-Moghaddam, M., & Ebrahimpour, R. (2019). Residual information of previous decision affects evidence accumulation in current decision. *Frontiers in Behavioral Neuroscience*, *13*, 1–12. https://doi.org/10.3389/fnbeh.2019.00009

Ostwald, D., Porcaro, C., Mayhew, S. D., & Bagshaw, A. P. (2012). EEG-fMRI based information theoretic characterization of the human perceptual decision system. *PLoS ONE*, *7*(4), e33896. https://doi.org/10.1371/journal.pone.0033896

Palestro, J. J., Bahg, G., Sederberg, P. B., Lu, Z. L., Steyvers, M., & Turner, B. M. (2018). A tutorial on joint models of neural and behavioral measures of cognition. *Journal of Mathematical Psychology*, *84*, 20–48. https://doi.org/10.1016/j.jmp.2018.03.003

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*(4), 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

Philiastides, M. G., Heekeren, H. R., & Sajda, P. (2014). Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. *The Journal of Neuroscience*, *34*(50), 16877–16889. https://doi.org/10.1523/JNEUROSCI.3012-14.2014

Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *The Journal of Neuroscience*, *26*(35), 8965–8975. https://doi.org/10.1523/JNEUROSCI.1655-06.2006

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Praamstra, P. (2006). Prior information of stimulus location: Effects on ERP measures of visual selection and response selection. *Brain Research*, *1072*(1), 153–160. https://doi.org/10.1016/j.brainres.2005.11.098

Praamstra, P., Boutsen, L., & Humphreys, G. W. (2005). Frontoparietal control of spatial attention and motor intention in human eeg. *Journal of Neurophysiology*, *94*(1), 764–774. https://doi.org/10.1152/jn.01052.2004

Praamstra, P., & Oostenveld, R. (2003). Attention and movement-related motor cortex activation: A high-density EEG study of spatial stimulus–response compatibility. *Cognitive Brain Research*, *16*(3), 309–322. https://doi.org/10.1016/S0926-6410(02)00286-0

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

Ratcliff, R. (2018). Decision making on spatially continuous scales. *Psychological Review*, *125*(6), 888–935. https://doi.org/10.1037/rev0000117

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision*, *2*(4), 237. https://doi.org/10.1037/dec0000030

Ratcliff, R., & McKoon, G. (2004). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922. https://doi.org/10.1037/rev0000117

Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences*, *106*(16), 6539–6544. https://doi.org/10.1073/pnas.0812589106

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., & Smith, P. (2015). *Modeling simple decisions and applications using a diffusion model*. Oxford University Press.

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

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Rihs, T. A., Michel, C. M., & Thut, G. (2007). Mechanisms of selective inhibition in visual spatial attention are indexed by?-band EEG synchronization. *European Journal of Neuroscience*, *25*(2), 603–610. https://doi.org/10.1111/j.1460-9568.2007.05278.x

Sagar, V., Sengupta, R., & Devarajan, S. (2019). Occurrence of the potent mutagens 2-nitrobenzanthrone and 3-nitrobenzanthrone in fine airborne particles. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-018-37186-2

Sewell, D. K., Jach, H. K., Boag, R. J., & Heer, C. A. V. (2019). Combining error-driven models of associative learning with evidence accumulation models of decision-making. *Psychonomic Bulletin & Review*, *26*(3), 868–893. https://doi.org/10.3758/s13423-019-01570-4

Shadlen, M. N., & Kiani, R. (2013). Decision making as a window on cognition. *Nature*, *80*(3), 791–806. https://doi.org/10.1016/j.neuron.2013.10.047

Smith, P. L. (2016). Diffusion theory of decision making in continuous report. *Psychological Review*, *123*(4), 425–451. https://doi.org/10.1037/rev0000023

Spiegelhalter, D. J., Carlin, B. P., & Linde, A. V. D. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, Series B (Statistical Methodology) *64*(4), 583–639. https://doi.org/10.1111/1467-9868.00353

Spilcke-Liss, J., Zhu, J., Gluth, S., Spezio, M., & Glascher, J. (2019). Semantic incongruency interferes with endogenous attention in cross-modal integration of semantically congruent objects. *Frontiers in Integrative Neuroscience*, *13*, 53. https://doi.org/10.3389/fnint.2019.00053

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*(3), 251–260. https://doi.org/10.1007/BF02289729

Tillman, G., Van Zandtc, T., & Loganb, G. D. (2020). Sequential sampling models without random between-trial variability: The racing diffusion model of speeded decision making. *Psychonomic Bulletin & Review*.

Turner, B. M., Forstmann, B. U., & Steyvers, M. (2019a). *Joint models of neural and behavioral data*. Springer.

Turner, B. M., Palestro, J. J., Miletíc, S., & Forstmann, B. U. (2019). Advances in techniques for imposing reciprocity in brain-behavior relations. *Neuroscience & Biobehavioral Reviews*, *102*, 327–336. https://doi.org/10.1016/j.neubiorev.2019.04.018

Turner, B. M., Rodriguez, C. A., Norcia, T. M., McClure, S. M., & Steyvers, M. (2016). Why more is better: Simultaneous modeling of EEG, fMRI, and behavioral data. *NeuroImage*, *128*, 96–115. https://doi.org/10.1016/j.neuroimage.2015.12.030

Turner, B. M., Wang, T., & Merkle, E. C. (2017). Factor analysis linking functions for simultaneously modeling neural and behavioral data. *NeuroImage*, *153*, 28–48. https://doi.org/10.1016/j.neuroimage.2017.03.044

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550. https://doi.org/10.1037/0033-295X.108.3.550

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, *111*(3), 757–769. https://doi.org/10.1037/0033-295X.111.3.757

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods*, *16*(1), 44–62. https://doi.org/10.1037/a0021765

Van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *The Journal of Neuroscience*, *28*(8), 1816–1823. https://doi.org/10.1523/JNEUROSCI.1853-07.2008

van Schouwenburg, M. R., Zanto, T. P., & Gazzaley, A. (2017). Spatial attention and the effects of frontoparietal alpha band stimulation. *Frontiers in Human Neuroscience*, *10*, 658. https://doi.org/10.3389/fnhum.2016.00658

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology. *Experimental Psychology*, *60*(6), 385. https://doi.org/10.1027/1618-3169/a000218

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Wagenmakers, E. J., Ratcliff, R., Gomez, P., & Iverson, G. J. (2004). Assessing model mimicry using the parametric bootstrap. *Journal of Mathematical Psychology*, *48*(1), 28–50. https://doi.org/10.1016/j.jmp.2003.11.004

Wagenmakers, E. J., Van Der Maas, H. L., & Grasman, R. P. (2007). An ez-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). Hddm: Hierarchical Bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, *7*, 14. https://doi.org/10.3389/fninf.2013.00014

Wilsch, A., Mercier, M. R., Obleser, J., Schroeder, C. E., & Haegens, S. (2020). Spatial attention and temporal expectation exert differential effects on visual and auditory discrimination. *Journal of Cognitive Neuroscience*, *32*(8), 1562–1576. https://doi.org/10.1162/jocn_a_01567

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, *39*(2), 293–306. https://doi.org/10.1016/S0042-6989(98)00114-X