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

Amin Ghaderi-Kangavari[1], Jamal Amani Rad[1,*], Kourosh Parand[2,3], and Reza Ebrahimpour[4,5]

[1]Department of Cognitive Modeling, Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran
[2]Faculty of Mathematical Sciences, Department of Data and Computer Sciences, Shahid Beheshti University, Tehran, Iran
[3]Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada
[4]Faculty of Computer Engineering, Shahid Rajaie Teacher Training University, Tehran, Iran
[5]Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[*]Corresponding. j_amanirad@sbu.ac.ir, j.amanirad@gmail.com

ABSTRACT

Model-based cognitive neuroscience consolidates the cognitive processes and neurophysiological oscillations which are reflections of behavioral performance (e.g., reaction times and accuracy). Here, based on one of well-known sequential sampling models (SSMs), named the diffusion decision model, and the nested model comparison, we explore the underlying latent process of spatial prioritization in perceptual decision processes, so that for estimating the model parameters (i.e. the drift rate, the boundary separation, and the non-decision time), a Bayesian hierarchical approach is considered, which allows inferences to be done simultaneously in the group and individual level. Moreover, well-established neural components of spatial attention which contributed to the latent process and behavioral performance in a visual face-car perceptual decision are detected based on the event-related potential (ERP) analysis. Our cognitive modeling analysis revealed that the non-decision time parameter provides a better fit to the top-down attention with the measures of two powerful weapons, i.e. the deviance information criterion called DIC score and R-square. Also, using multiple regression analysis on the contralateral minus neutral N2 sub-component (N2nc) at central electrodes and contralateral minus neutral alpha power (Anc) at posterior-occipital electrodes in the voluntary attention, it can be concluded that poststimulus N2nc can predict reaction time (RT) and non-decision time parameter relating to spatial prioritization. Whereas, the poststimulus Anc only can predict the RT and not the non-decision time relating to spatial prioritization. The result suggested that the difference of contralateral minus neutral oscillations was more important to reflect the modulation of the top-down spatial attention mechanism in comparison with the difference of ipsilateral minus neutral oscillations.

Introduction

Decision-making is a high-level cognitive process whereby the decision-maker makes different decisions based on the evidence available, expected value and the possible outcomes ([1]). This process, the decision, operates in a flexible time frame so that it is not in the form of immediate evidence acquisition and the real-time demands of the action itself ([2,3]). The most frequent type of decision that we routinely make in different situations is perceptual decision making, a process that involves the processing of visual information, attention, the accumulation of evidence and motor response ([4]). Hence, many perceptual decisions might bring about disastrous consequences by choosing the wrong choice. Suppose that you are heading 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 (coherency), then response time and accuracy are affected. For example, if the weather is rainy and dark, we cannot effortlessly detect the traffic light and therefore our performance will fall down. In addition, when a driver is moving toward the intersection in rainy and dark weather, she/he do not necessarily know the exact spatial location of the traffic light. Moreover, at first, the driver should struggle to locate the locus of the traffic light (stimulus), then process the received information from the light. If the driver is told exactly where the light is (spatial attention) before reaching the intersection, the driver will use the focus of his attention to the target and suppress attention to others to detect areas faster and perhaps even more accurately. ([3,5,6]). Therefore, it can be said that attention plays a key role in the perceptual decision-making process ([7–10]). Besides, Posner ([11]) studies have shown that people can pay close attention

without moving their eyes to an area of space. Moreover, one type of spatial attention is goal-driven, pointing to top-down or endogenous attention that is a situation that we attend to a region of space where the stimulus is present (valid) or not present (invalid) ([11]). The top-down cue could facilitate behavior performance and manipulate neural mechanisms ([12,13]) which some aspects are still unknown. Therefore, in this paper, we have just concentrated on valid covert attention on distracted stimuli which were presented by endogenous orienting cues.

Although perceptual decision making involves a lot of components, two aspects whose interactions have hardly received enough notice are coherency and spatial attention, which can deeply affect the performance of individuals. Therefore, this research strives to elaborate biomarkers of spatial prioritization in the electroencephalogram (EEG) to predict behavioral performance and a well-established model's components in a face-car perceptual decision making task under task difficulty. To be more specific, we designate whether the lateralized alpha frequency band at parietal-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 behavioral performance and parameters of spatial prioritization in perceptual decision making ([7,14]).

Top-down spatial attention could modulate alpha band (8-13 Hz) of down stream area such as posterior-occipital sites to direct and allocate limited processing resources to attended and unattended loci ([15,16]). In fact, studies shows that the increasing alpha frequencies ipsialaterally (alpha-synchronization) band relating to unattended location could suppress task-irrelevant space, and the decreasing alpha frequencies contralaterally (alpha-desynchronization) band related to attended location intentionally facilitate upcoming visual processing ([15,17]). This inverse mechanism of top-down orienting of attention is found usually after postcue (prestimulus) of various tasks such as visual ([18,19]), tactile ([20]), and auditory ([21]). Another important factor which has the substantial role to prediction distributed finite resources for spatial prioritization in central electrode sites is central contralateral N2 (N2cc) subcomponent ([22]). The N2cc deflection is activated by motor selection and preparation processes of particular space location which affects to the response time (RT) measured within the motor execution of the task ([23]). The N2cc amplitude which mirrors motor preparation mechanism is contracted with informative top-down cue ([22]). Moreover, researchers have reported that the N2cc deflected over central electrodes sites might coincide in time with the N2 posterior contralateral (N2pc), but the important finding of N2cc is not due to overlap of the N2pc with movement execution-related activity ([14,22]). The other hand, the anterior contralateral N2 subcomponent (N2ac) appears in order to regulation sound localization performance[7,24]) and also emerges in the visual Simon task and ([25]).

The goal of this study is to find relationship between these two intriguing biomarkers, lateralized alpha frequency band over parietal-occipital electrode sites and N2-subcomponent amplitude, with behavioral performance and the non-accumulation time, also called the non-decision time ([26]), in mathematical modeling which has been designated as the best parameter by fitting performance data via a powerful nested model comparisons approach ([27]).

0.1 Cognitive modeling of perceptual decisions

Perceptual decision making is a process, so identifying the components of this process is very worthwhile. Hence, the purpose of behavioral modeling is to understand what components or parameters can be extracted from decision making process and how different people may commit mistake errors or correct response ([28]). There is a growing consensus that the evidence accumulates gradually and sequentially to make a decision ([29,30]). As a result, sequential sampling models (SSMs; Stone, 1960 ([31]), Ratcliff, Smith, Brown, and McKoon, 2016 ([32]) and Evans and Wagenmakers, 2020 ([33]) for the reviews), as the most well-known explanation of how the decision-making process works ([26,30]), have obtained very achievements in modeling the cognitive processes underlying decision making across a wide variety of paradigms, such as optimality polices ([34–37]), stop signal paradigms ([38]), go/no-go paradigms ([39]), multi-attribute & many alternatives choice ([40–42]), learning strategies ([43–45]), attentional choice ([8–10,46]), continuous responses ([29,47]), neural processes ([1]), and so on. In general, SSMs assume that decisions are made from a noisy process of accumulating evidence, that is to say, according to this theory, the evidence is gradually accumulated in favor of different choice alternatives over time with some rate until a sufficient amount of evidence for one of the options reaches a predetermined threshold to make a decision. The DDM which has been growing in perceptual decision-making modeling operates on a random process (Wiener or Brownian-based process) because human decision-making seems to be a random process which means that in the same condition, the individual may respond differently to the same stimulus ([32,48]). In other words, the model can be considered as a bridge between behavioral data and cognitive processes. Hitherto, various sequential sampling models have been presented by the researchers. There are some successful sequential sampling models such as: EZ-diffusion model ([49]), linear ballistic accumulators ([50]), leaky-competing accumulator ([51]), and race diffusion model ([52]) but perhaps, the drift-diffusion model (DDM) is the most popular model that being utilized in information processing modeling ([30]). 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, and by recalling a piece of evidence, the accumulator steps toward the upper/lower boundary, and by hitting one of these boundaries the accumulation process is stopped and the corresponding option is selected

([53]), illustrated in Figure 1. Mathematically, evidence accumulation in the DDM can be described continuously as,

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

or 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 (as shown in Figure 1), 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 lays on zero and the upper boundary has distance 'a' from the lower boundary. Therefore, 'a' shows the threshold parameter, and actually is an index for the speed-accuracy tradeoff. It means by higher values of 'a', the participant should accumulate much amount of information to make a decision and his/her decision is more accurate but it takes too long. On the other hand, by lower values of 'a', the participant makes the decision faster but not accurate enough ([35]). The second parameter of the DDM is the drift rate which is denoted by 'v'. The drift rate is the mean slope of fluctuation of the accumulator toward the boundaries. This parameter shows the speed of information processing and identifies task difficulty level. In fact, higher values of the drift rate show that the task has no demanding cognitive load and it is easy to make a decision between the options. On the other side, when the value of the drift rate tends to zero, shows that the task is demanding and the speed of information processing is slow ([54]). The third parameter is the bias which is shown by 'zr' and its implication is the starting point of the accumulation process. When the starting point is equal to '$\frac{a}{2}$', the distance between the starting point and the boundaries are equal and there is no bias to each option. But when 'zr' is greater or less than '$\frac{a}{2}$', then the participant has a bias to upper or lower boundaries, respectively ([55]). 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 reaction time. Usually, this parameter is denoted by T_{er} or 't_0' in DDM ([56]). These are the main parameters of the DDM and there are some between trial variability parameters in addition. s_v, s_{zr}, and s_t are the between trial variability parameters for drift rate, bias, and non-decision time respectively. Usually, normal distribution is considered for s_v and uniform distribution for s_{zr} and s_t ([9,26]). High values for the between trial variability parameters indicate that there is some variability in materials or in cognitive states ([57]).

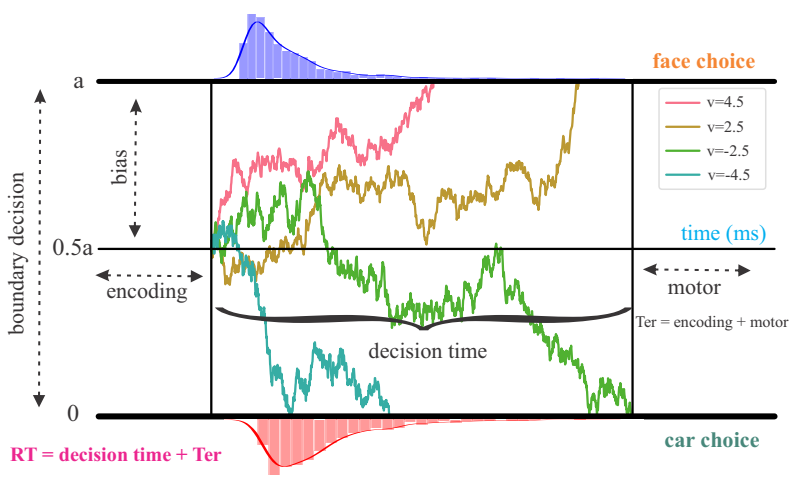


Figure 1. Trajectories of drift-diffusion model for two stimuli with various drift rates.

To come to the point, we hypothesize that electrophysiological correlates of top-down spatial prioritization (alpha band lateralization and N2 subcomponent) during a visual perceptual decision making task could predict visual performance and the non-decision time parameter of DDM model reflecting the sensory coding time and response expectation. This claim is based on that the cognitive processes mirrored by lateralized alpha power in around occipital electrode sites and lateralized N2cc amplitudes in central electrode sites are modulated by motor response area and decoding stimulus area could regulate to choose the distracted visual face and car stimuli, so they should contribute to the non-decision processes and the performance.

1 Methods

1.0.1 Experiment overview

We re-examine the data from an experiment conducted by ([58]), on visual coherence and spatial attention effects in perceptual decision-making. In their study, seventeen participants (8 females, mean age 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 reimbursed in various ways. Electroencephalograms (EEG) and behavioral data in perceptual decision making task under spatial prioritization were acquired over two data acquisition sessions. In the task, each of two independent random variables, coherency and spatial attention, were used so that each variable had two levels. In fact, the independent variable of the coherency had two level "low" and "high", and the independent variable of the spatial attention had also two level "yes" and "no". For more details about experimental design & stimuli, interested readers can refer to the main report ([58]) written on the data by Georgie et al.

Also, their study design was a two-way repeated-measure factorial design (please see Figures 2a and 2b). As you can see in Figures 2a and 2b, in each trial, in order to manifest the informative and uninformative spatial prioritization, the cuing one-way arrow and the cuing two-ways arrow respectively are informed for one second period before the stimulus emergence. Then, a visual stimulus from the car or face is displayed to the subject for 200 milliseconds, and therefore subjects are asked to press a button to respond whether they perceived the car or the face. Moreover, the coherency of the stimuli is regulated. Subjects should use the index finger and the middle finger of the right hand to respond and look at the fixation point even during all the task period, so that they should attend on one side (informative prioritization) or two sided (uninformative prioritization) covertly. Therefore, in half of the trials, the one-way arrow is used to indicate the presence of the stimulus in one of our field of vision, and in another half of the trials, a two-way arrow is used to indicate that the presence of the stimulus is not reported and subject must pay attention to both fields of vision. The task procedure is as follows

Following ([58]), for each of the conditions (coherency, spatial attention, and stimulus), 36 trials for both EEG and behavioral data had been acquired with an inter-trial period between 0 ms and 300 ms as well as a sampling rate of 5000 Hz points. Also, all trials from each condition had been randomly selected to participants. Besides, there were 64 EEG sensors consisted of 62 scalp electrodes via the 10-20 system and two extra sensors that one was designed roughly 2 cm under the left collarbone in an effort to identifying electrocardiogram (ECG), and another one was designed under the left eye in an effort to the recording of eye-blinking artefacts (ECG).

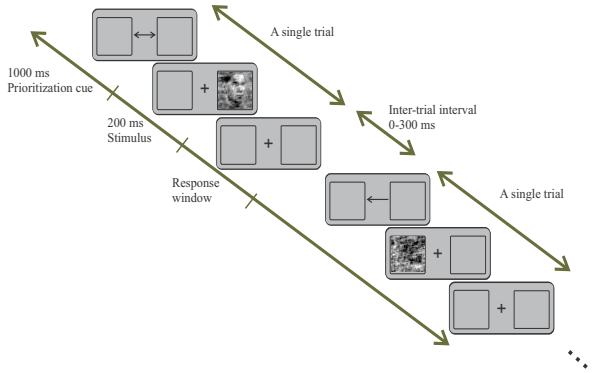
In this paper, we exclude three participants because their EEG data are very contaminated and not able to be clean, so fourteen participants enter the further analysis. Also, in order to statistical analysis and model fitting, we remove outliers behavioral data using a kind of the IQR Interquartile range which IQR = Q3 - Q1 as the distance from the first quartile to the third quartile. To compute the interquartile range for the data, any number smaller than Q1 - 3 $\times$ IQR and greater than Q3 + 3 $\times$ IQR are suspected outlier. Consequently, reaction time and accuracy are submitted to parametric 2 $\times$ 2 $\times$ 2 repeated-measure analysis of variance (ANOVA) statistical test.

1.0.2 EEG signal analysis

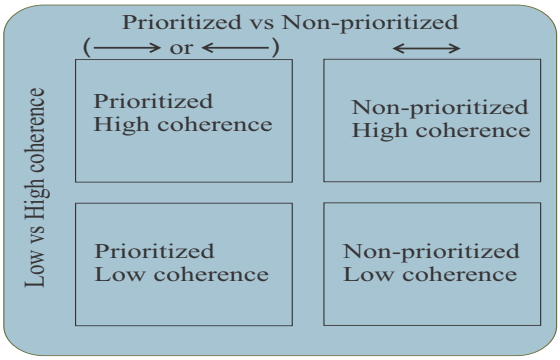
As well mentioned in ([58]), the EEG data were recorded using a 64-channel electroencephalograms system at a sampling rate of 5000 Hz samples. As we know, the EEG data routinely record as the continuous (Raw) signal which contains of the real source of brain signal, biological or non-biological artifact and noise ([59]). Therefore, in this work, the event-related potential (ERP) approach which prepares a worthwhile technique to detect cognitive components across conditions or between groups in an experiment is used. Even though a lot of spatial attention studies have focused on pre-stimulus amplitude components and alpha oscillation, in this study we are just interested to concentrate on poststimulus components in the perceptual decision making task under spatial attention. The EEG preprocessing and analysis were implemented by the MNE-python package ([60]). In fact, the preprocessing steps we have done can be summarized as follows: the continuous EEG data were down-sampled to 256 points, band-pass filter in the range 1 Hz - 30 Hz, re-referencing with average of sensors, visual inspection to remove abnormal frequency, interpolation fully corrupted signal, discarding range of -100 mV to +100 mV amplitude, extracting Epochs from - 100 ms (baseline) of cue's onset to 600 ms after stimulus appearance (totally 1700 ms), running Independent Component Analysis (ICA) to remove manually irrelevant-task (e.g. eye movement, head motion and muscular activity), detecting automatically eye-blinking EOG component and heartbeat peak ECG component. The EEG preprocessing with MNE-package implementation code in python named 'MNE-preprocessing' is readily accessible at https://github.com/AGhaderi/MNE-Preprocessing.

The time-Frequency approach decomposes multiresolution properties of non-stationary EEG signal that comprises both the phase-locked and the non-phase-locked oscillations (trial-to-trial variability) ([59,61]). Therefore, to explore the spectral content of the data, we have employed the Morlet wavelets transform analysis which is an efficient time-frequency decomposition over every single trial. Consequently, after performing single-trial Wavelet decomposition from 4Hz to 30Hz, the representation has been averaged across trials for each subject. To setting the parameter of "number of cycles" which controls the trade-off between temporal precision and frequency precision, we adjusted three-cycle wavelets at the lowest frequency (4 Hz) and 16-cycle wavelets at the highest frequency (30 Hz) by increasing the rate of 0.5.

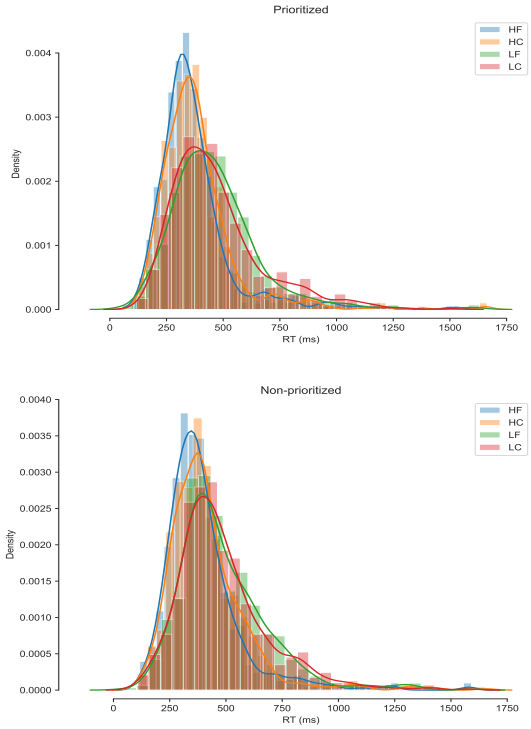
In this paper, the N2 subcomponent Amplitude and Alpha band frequency are explored to be fed into a regression model. Therefore, we calculated the mean contralateral and ipsilateral (informative) and neutral evoked potentials (uninformative) at C1/2/3/4 central electrodes, and also the mean ipsilateral and contralateral power across time in the alpha frequency band (8–13



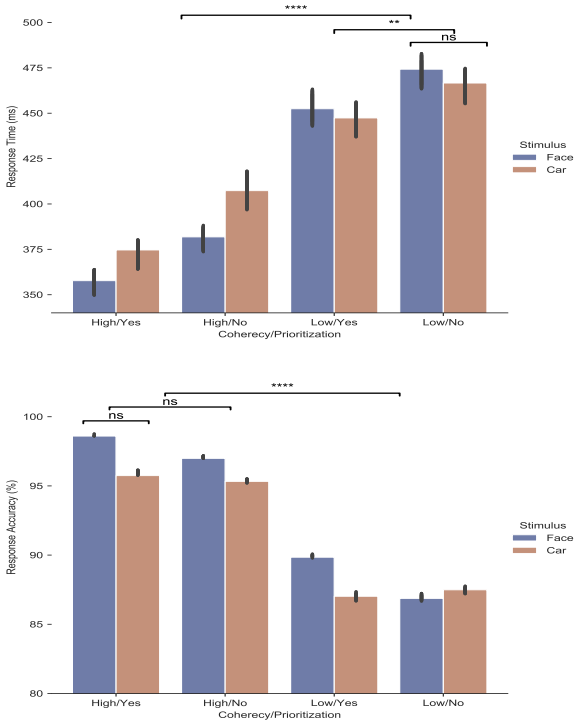
(a) Task procedure.



(b) The 2×2 study design.



(c) Behavioral data (RT distributions)



(d) Behavioral data (RT and accuracy)

Figure 2. Experimental design and behavioral data. (A-B) Overview of experiment task and the 2×2 study design. (C) Reaction time (RT) distributions of face and car for all four conditions (two coherences and two prioritizations). HF: high coherence and face stimulus, HC: high coherence and car stimulus, LF: low coherence and face stimulus, LC: low coherence and car stimulus. (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 'np' indicates non-significant p value.)

Hz) at parietal-occipital electrodes PO3/4/7/8 ([14,22,23]). The contralateral response contains the average ERP oscillation at left hemispheric sensors in right-way cue and right hemispheric sensors in left-way trials, the ipsilateral response contains the average ERP oscillation at left hemispheric sensors in left-way cue trials and right hemispheric sensors in right-arrow trials, and eventually the neutral response involves the average ERP oscillation at left and right hemispheric sites in two-way cue trials. Therefore, the previous researchers have been focused on prestimulus alpha oscillations and N2 subcomponent to predict spatial prioritization, but in this study, we have concentrated on the poststimulus of these regressors as the anticipatory indexes for allocating of spatial prioritization. In an effort to measure the N2 subcomponent, the dependent sample t-test was computed between averaging contralateral and neutral amplitude (N2nc), and between mean ipsilateral and neutral amplitude (N2ni) at a time window within 150-500 msec (90 time points) after stimulus onset (or 1150-1500 msec after cue onset). This signifies that there are 90 statistical t-tests, so the False Discovery Rate (FDR) correction procedure which gives us a different standard for significance was utilized to correct the t-tests ([59]). Afterward, the measurement window of N2cc and N2ci was based on significant time points by FDR correction. On the other hand, to assess alpha lateralization, the contralateral-neutral alpha band (Anc) index, and the ipislateral-neutral alpha band (Ani) index were conducted across both hemispheres as follow:

Anc = (contralateral alpha power - neutral alpha power),

Ani = (ipsilateral alpha power - neutral alpha power).

When the Anc index is negative, the contralateral alpha power is lower (red color) over the neutral alpha band at parietal-occipital sensors and vice versa. In contrast, when the Ani index is positive values, the ipsilateral alpha power is higher (blue color) over the neutral alpha band and vice versa. As a result, to determine the measurement window of alpha power, a 200-msec time window was applied around the 50% Fractional Area Latency (FAL) to Estimate the Midpoint Latency (or local peak) in the grand-averaged ANC and ANI indexes ([7,59]). Accordingly, we computed 50% FAL contingent on a wide time window within 0-600 after stimulus onset (or 1000-1600 after cue onset). Therefore, each couple of variables of {N2nc, Anc}, and {N2ni, Ani} played the role of regressors in order to the regression model.

1.0.3 Cognitive modeling and model comparison

In an effort to explore latent cognitive processes underlying perceptual decisions under spatial prioritization, and link them to neural mechanisms, we use a hierarchical Bayesian estimation approach. The two main assumptions in the Bayesian hierarchical approach are that, first, the participant-specific parameter values are randomly sampled from a group-level distribution that is updated using data from all participants, and second, that the distribution of the parameters is estimated simultaneously at the group level and the individual level ([62]). Therefore, combining this approach with the diffusion model results in a stable estimation of parameters (i.e. drift rate, the boundary separation, and the non-decision time) at both individual and group levels. In this work, this model-fitting approach is done using the HDDM package in Python ([63]) which generates full posterior distributions of the model parameters at the individual and group level by the Markov-chain Monte Carlo sampling (MCMC). This package is flexible and the user-friendly program produces more stable results and appropriate parameter estimation in the form of posterior distributions for each parameter included in the model when less data (i.e., experimental trials) are available compared to other implementations such as traditional maximum likelihood estimation for individual subjects (for details about the model see[63,64]).

The HDDM package offers a model parameterization depending on the experimental factors, e.g., one could model different drift rates for all conditions or for all groups or any combination of them[65]. Different criteria such as the deviance information criterion (DIC) can be used for a fair comparison of these different types of Bayesian models. On the other hands, in the full DDM, there are seven parameters which might be categorized into three parts: the decision process measures (boundary decision, mean starting point, and mean drift rate), the non-decision process measure (non-decision time), the inter-trial variability measurement (drift rate variability s_v, starting point variability s_z, and non-decision time variability s_t) ([32,48]). Based on the strong evidence previously presented in various studies to investigate the effect of coherency as well as stimulus on the decision-making process ([4,27,66]), we propose cognitive models whose drift rate parameter is coherence+stimulus dependent. Therefore, in order to the suggested hypothesis in the current paper, spatial attention regulated resources of decision making, there are five possible nested model procedures to detect how spatial attention affect the decision making procedure ([27]). First, it may do not capture or shift any resources or decision parameters, so it could be the pure model with no parametric dependence on spatial prioritization (we call this model $model_p$, in which only the drift rate parameter depends on coherence+stimulus). Second, the spatial prioritization could regulate the rate of accumulated evidence (the drift rate v) named $model_v$. Third, it is free to vary with the non-decision time (T_{er}) named $model_t$. Four, it manipulate the mean bias (starting point z) named $model_z$. Five, it might shifts the boundary decision (a) called $model_a$. Additionally, the parameter of drift rate was free on the two level of the coherency (high and low), in line with the related finding about stimulus strength in perceptual tasks ([4,67]). Moreover, inter-trial variability non-decision time 'st' for group level was free in all models, wherease other inter-trial variablities were not estimated. Then, there are two ways to implementation of DDM: an accuracy-based approach which correct responses are coded 1 and incorrect responses 0, or an stimulus-based approach which response stimuli as face and car are measured. In this work, the stimulus-based approach is employed.

For each model, Markov Chain Monte Carlo simulations ([68]) were used to generate 100,000 samples from the joint posterior distribution of parameters using the HDDM package ([63]), from which the initial 1000 samples were discarded as a burn-in phase to minimize the effect of initial values on the posterior inference and converge on stable distributions ([63]). The convergence of the Markov chains was assessed through visual inspection as well as by calculating the R-hat Gelman-Rubin statistic ([69]) to ensure that the models had properly converged, which compares between-chain and within-chain variance. All R-hat values were near 1.0 (within 0.0001) for the parameters, indicating that convergence had been achieved ([70]). Also, both the R-square and the deviance information criterion (DIC) of each model which evaluate the model's goodness-of-fit while accounting for model complexity (i.e., number of free parameters), with lower DIC values indicating better model fit[71], have been used in order to model comparison method to choose the best scenario of the diffusion model in group level ([18,63]).

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 ([63]), see supplementary material. The hDDM implementation code and simulation scripts for all current models are readily available at https://github.com/AGhaderi/hDDM_attention.

1.0.4 Multiple regression model

In this section, we explored whether both couple explanatory (independent) variables of {N2nc, Anc}, and {N2ni, Ani} could predict spatial prioritization of the best scenarios of four scenarios in the previous section and behavioral performance (RT and accuracy). Then, the multiple regressions using the Ordinary Least Squares (OLS) method have been applied by a statistics module (statsmodels) in Python as the following formula:

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

where two situations happened for couple regressors, X_1 is $N2nc$ and X_2 is Anc, or X_1 will be another option $N2ni$, and X_2 will also be another option Ani. In two different situations, y is the parameter behavioral performance (RT or accuracy) which leads to four multiple regressions. In addition, the multiple regression model is able to assess the relationship between electrophysiological correlates and latent processing of spatial prioritization resulting from modeling parameters. As a result, p-values for each beta coefficient and R-squared is reported as a goodness-of-fit statistic that displays the relationship between the observed and predicted values. Component-Component plus Residual (CCPR) or the partial Residual plot (graphics.plot_ccpr function from statsmodels module) is a worthwhile metric that can be utilized to assess the impact of a regressor on the outcome variable (y) by taking into account the effects of the other independent variables which is formulated as $\beta_i X_i + \varepsilon$ versus X_i. The CCPR plot provides $(\beta_i X_i, X_i)$ line to manifest where are points, so that you can see whether the relationship between the variation in a regressor on response variable seems to be linear or not.

2 Results

2.0.1 Behavioral result

Both coherency and spatial prioritization were manipulated at two distinct factors which changed the quality of sensory information and the top-down intention relating to a specific visual field respectively. At each trial of the face versus car task, the subjects could respond as quickly and accurately as to the emerged stimuli. However, in order to testify of the experimental manipulation, the behavioral performance was analyzed to assess the effect of these variables, so here is the result of three-way repeated measure ANOVA. For response time, this ANOVA found no significant main effects of stimulus ($F(2,13) = 0.78$, $p = 0.39$), but revealed significant main effects of spatial prioritization ($F(2,13) = 11.83$, $p < 0.01$) and phase coherence ($F(2,13) = 38.28$, $p < 0.0001$) (Fig. 2d). Therefore, the rise of the phase coherence and spatial prioritization versus spatial non-prioritization led to a more accurate response and faster reaction time. Interaction between stimulus and coherency was also significant ($F(2,13) = 8.54$, $p < 0.05$), but other interaction between variables were not statistically significant, i.e. interaction of spatial prioritization and phase coherence ($F(2,13) = 0.54$, $p = 0.47$) and interaction of spatial prioritization and stimulus ($F(2,13) = 0.13$, $p = 0.71$). Because of the significant interaction between coherence and stimulus, it is worthwhile to create a model relating to stimulus coding, not just precise coding, so that they are free in a common parameter (drift rate). Finally, for statistical analysis of accuracy, the coherency had only a significant main effect ($F(2,13) = 41.83$, $p < 0.0001$) (Fig. 2d). The main effect of stimulus ($F(2,13) = 0.84$, $p = 0.37$) and spatial prioritization ($F(2,13) = 2.12$, $p = 0.17$) was not significant across their factors, i.e. interaction of spatial prioritization and phase coherence ($F(2,13) = 0.05$,$p = 0.81$), interaction of spatial prioritization and stimulus ($F(2,13) = 2.9$, $p = 0.11$) and interaction of coherence and stimulus ($F(2,13) = 0.14$, $p = 0.71$).

2.0.2 Cognitive modeling

In an effort to explore the underlying mechanism of spatial prioritization in perceptual decision making, we applied the well-established hierarchical Bayesian drift-diffusion (HDDM). Freedom of models parameters could create opportunities to investigate the effect of spatial prioritization in the categorized face versus car task. However, five possible models were run to

give us appropriate intuition of spatial prioritization. All convergence evaluations of the models were performed with full and accurate details, which are reported in the supplementary material of the article. In instance the MCMC sampling traces of the group posteriors distribution for these parameters which show converged chain procedure during parameters estimation were utilized (see Figures 8, 9 and 10). Also, the convergence of the Markov chains was assessed by calculating the R-hat Gelman-Rubin statistic ([69]) to ensure that model$_t$ showed superior convergence and the same stationary distribution with four chains, based on R-hat values under 1.0001 (see supplementary material). Also, we compared posterior model predictions with the observed data, which was a good 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 reaction times across the four different conditions (see Figure 11). On the other hand, to compare which model is better to fit observed data, our analysis was equipped with two powerful measures i.e. DIC and R-square. the result of model comparison analysis reveals that model$_t$ provided the best model fit so that it received the smallest DIC and largest R-square, so non-decision time parameter (T_{er}) interpret perfectly the manipulation of spatial prioritization (see Table 3).

Here, we utilized the DDM to examine our hypotheses and identify the electrophysiological mechanisms of informative (contralateral and ipsilateral) and uninformative (neutral) prioritization under decision-making. Therefore, the alternative five models were run to determine which parameter and cognitive process were affected by two factors of top-down spatial attention. As a result, in order to assess parameters of the winner model$_t$ across conditions, we used the dependent Student's t-test for non-decision time and two-way repeated measure ANOVA for the drift-rate parameter. Non-decision processes of spatial prioritization (Yes) and non spatial prioritisation (No) were significantly different ($t = 3.34$, $p = 0.0053 < 0.01$), see Fig. 3a. In fact, non-decision processes take less time when people receive a cue of prioritization than when they receive no prioritization cue. In addition, the main effect of drift rate relating to coherency were significant ($F(1,13) = 75.11$, $P < 0.001$), and the drift rate relating to stimulus had no significant main effect ($F(1,13) = 0.3$, $P = 0.87$), see Fig. 3b. There were also no significant of interaction between stimulus and coherency of drift rate ($F(1,13) = 0.64$, $P = 0.43$). Finally, box-plots of non-decision time T_{er} and drift rate parameters in different conditions are reported in Figs. 3c and 3d, respectively.

Correlation of the DDM and conventional outcome Parameters: To improve the understanding, which characteristics of behavioral outcome are indicative of cognitive parameters, mean posterior values for non-decision time from model$_t$ affected on response times relating to spatial attention. Matrix correlations are visualized in Figure 4. As a result, $RT_{er,P-NP}$ (response time of prioritization minus response time of non-prioritization) and $T_{er,P-NP}$ (non-decision time of prioritization minus non-decision time of non-prioritization) are highly positively correlated '0.47', indicating that larger $T_{er,P-NP}$ in the decision process are associated with faster responses $RT_{er,P-NP}$.

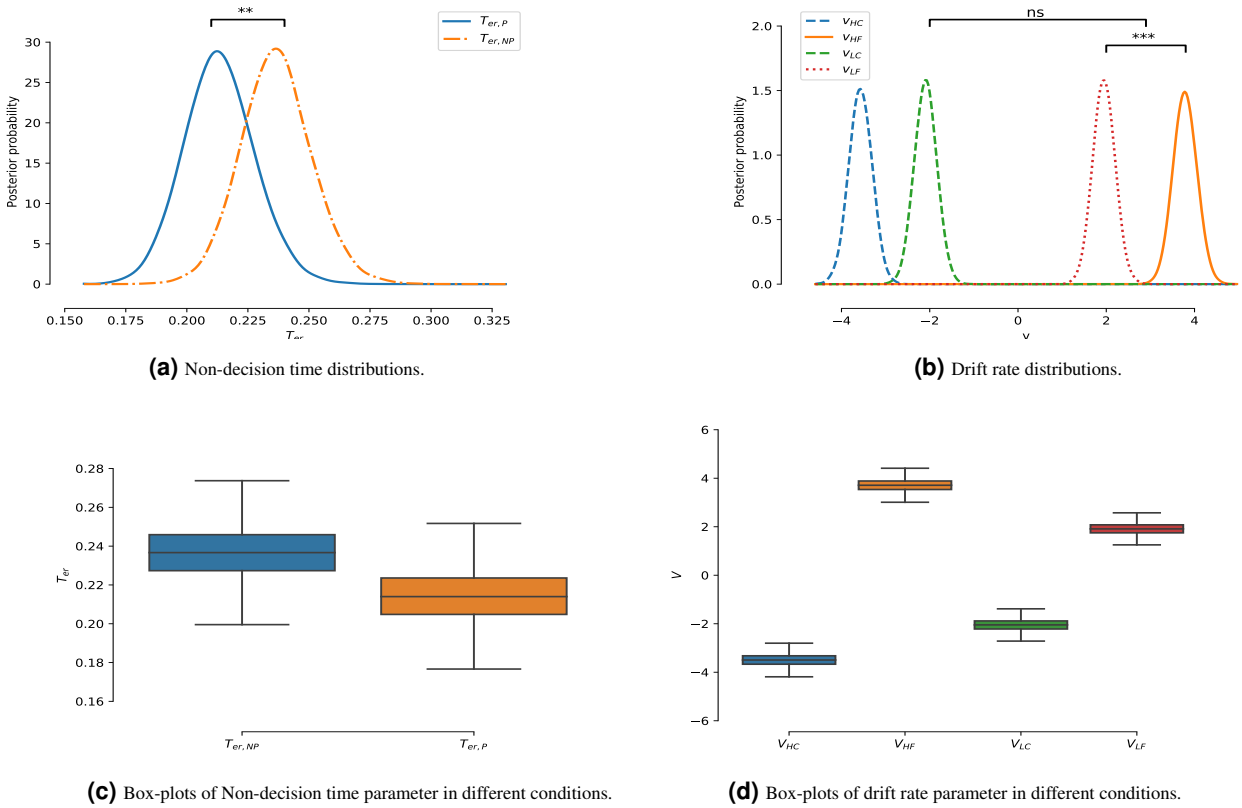
2.0.3 Central N2 subcomponent and alpha lateralization

The ERPs lateralization at central electrode C1/2/3/4 have been shown in Figure 5. Moreover, N2nc and N2ni which are displayed by the red color as columns were significant based on FDR correction applying statistical dependent t-test on the time window after stimulus onset. Figure 11e exhibits the grand-average of signal contralateral and neutral as well as the difference between them as time-locked to cue and stimulus array onset. FDR correction over samples of the time window found two distinct N2nc subcomponent at central site 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). Figure 11g presents the grand-average of signal ipsilateral and neutral and also ipsilateral minus neutral. As a result, the FDR correction over samples of the time window revealed an N2ni subcomponent at the central site with (mean $p = 0.022 < 0.05$ for time window 210-270 msec).

On the other hand, Figure 6 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). In addition, at first, we considered the ± 50 msec around of 50% FAL (237-337 msec) on -100 msec to 600 msec in Anc signal. Then, by applying a one-sample t-test in the mean portion of Anc, we had a significant effect with ($t = 2.59$, $p = 0.022 < 0.05$) which are displayed by the red color, see Figure 6a. Also, we examined the same analysis in the Ani signal which a time window (155- 255 msec) was concluded. 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 are shown by the gray color, see Figure 6a.

2.0.4 Regression analysis

In this section, in order to investigate the mechanism of spatial attention based on electrophysiological components in perceptual decision making, we applied the separate multiple regression model with prioritization minus non-prioritization of non-decision time parameter in DDM model and of behavioral performance as output and lateralization neutral activity of amplitude and alpha. Tables 1 and 2 show the result of model fitting for both of contralateral versus neutral component and ipsilateral versus neutral component as the difference between informative (contralateral or ipsilateral) versus uninformative (neutral) neutral activity. For reaction time of contralateral versus neutral, difference RTs were perfectly predicted by N2 sub-component contralateral minus neutral (N2nc) with time measurement 260-315 msec and alpha power contralateral minus neutral (Anc) with the time measurement 237-337 msec (R-square=0.679,$F(2,13)=11.66$, $p <0.01$). There was no significant effect for N2nc



(a) Non-decision time distributions.

(b) Drift rate distributions.

(c) Box-plots of Non-decision time parameter in different conditions.

(d) Box-plots of drift rate parameter in different conditions.

Figure 3. Posterior distribution of the group parameter non-decision time and drift rate from the model$_t$. (a), Non-decision two levels of spatial attention were significant different by dependent t-test (t = 3.34, p < 0.01). (b) The main effect of coherency were significant (F(1,13) = 75.11, P < 0.001), the main effect of stimulus were not significant (F(1,13) = 0.3, P = 0.87). There were also no significant of interaction between stimulus and coherency of drift rate (F(1,13) = 0.64 , P = 0.43). $T_{er,P}$: T_{er} of prioritization, $T_{er,NP}$: T_{er} of non-prioritization, V_{HC}: V of high coherence and car stimulus, V_{HF}: V of high coherence and face stimulus, V_{LC}: V of low coherence and car stimulus, V_{LF}: V 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.

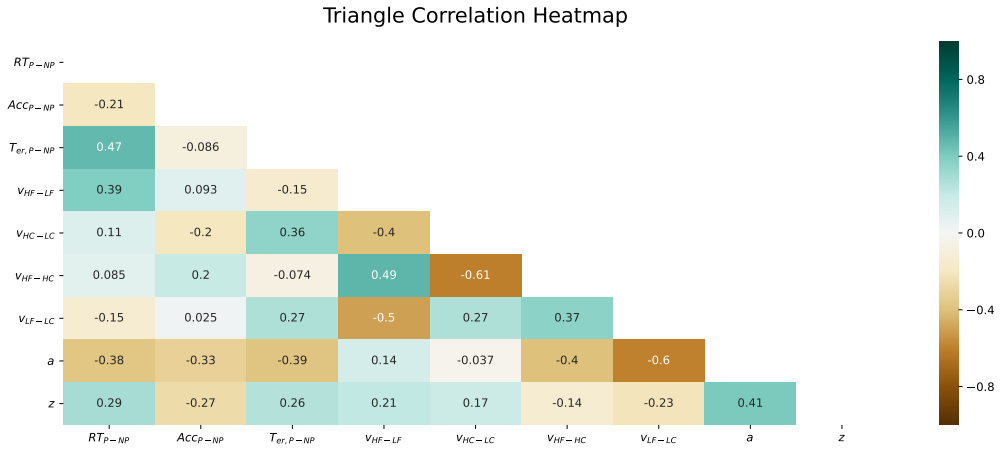
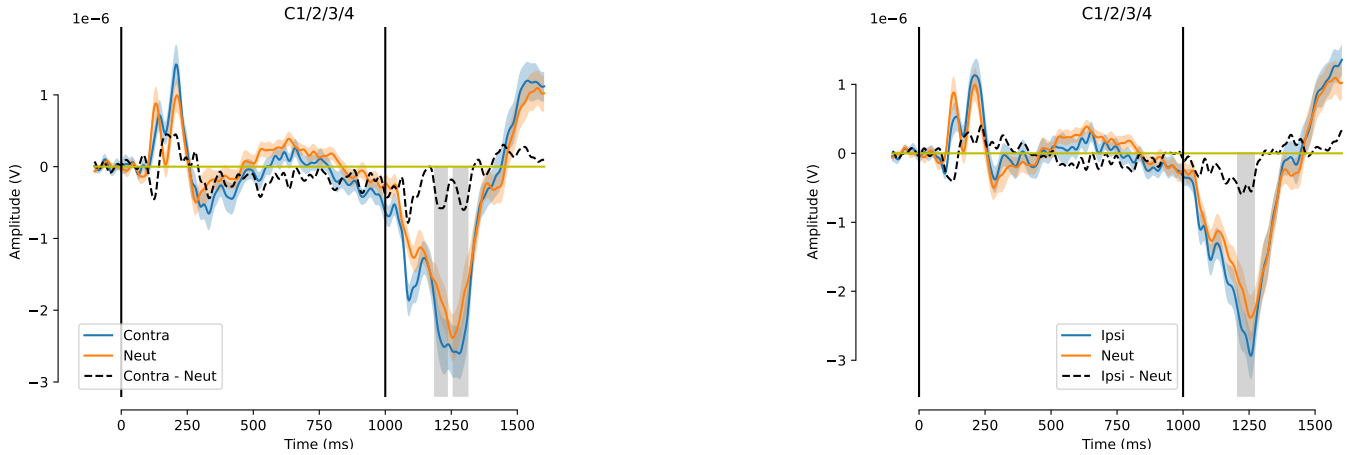


Figure 4. The heatmap matrix correlation of the group-level parameters across conditions. RT_{P-Np}: RT of prioritization minus RT of Non-prioritization, Acc_{P-Np}: Accuracy of prioritization minus Accuracy of Non-prioritization, $T_{er,P-Np}$: $T_{er,P}$ minus $T_{er,NP}$, , V_{HF-LF}: V_{HF} minus 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}



(a) Contralateral versus neutral ERP at central electrodes.

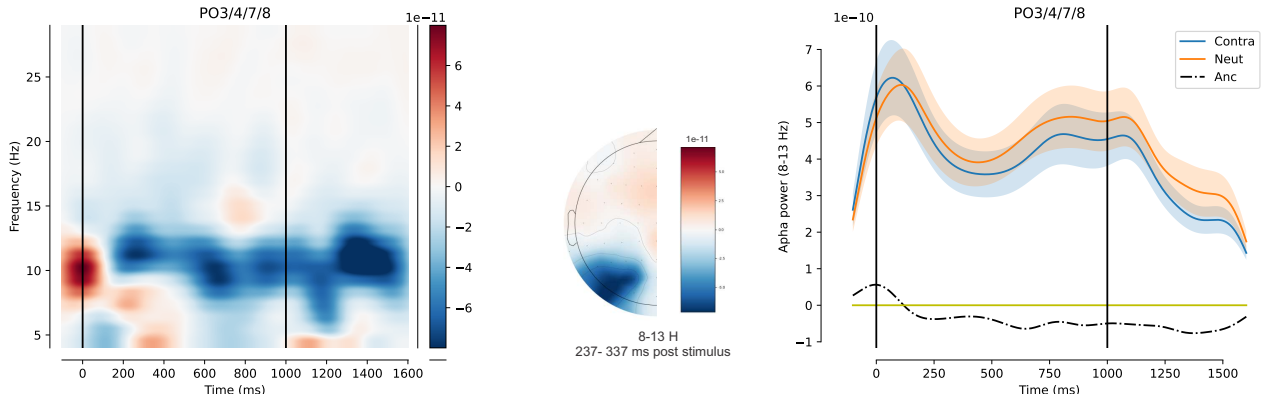
(b) Ipsilateral versus neutral ERP at central electrodes.

Figure 5. N2nc and N2ni subcomponents at central electrodes C1/2/3/4. (a) The grand-average contralateral and neutral amplitude of the ERP signal as well as the outcome of (contralateral minus neutral) are displayed. Moreover, the significant measurement window based on FDR correction of t-test with $\alpha = 0.05$ (the mean $p < 0.05$ for 190-235 msec and the mean $p < 0.05$ for 260-315 msec windows) on a broad time window from 150 msec to 500 msec after stimulus onset are illustrated by the red color, (b) The grand-average ipsilateral and neutral amplitude of the ERP signal as well as the outcome of (ipsilateral minus neutral) are displayed. In addition, the significant measurement window based on FDR correction (mean $p < 0.05$ for 210-270 msec time window) are illustrated by the gray color. Moreover, the first and second vertical lines show cue prioritization and stimulus presence respectively, and shaded regions identify the standard error.

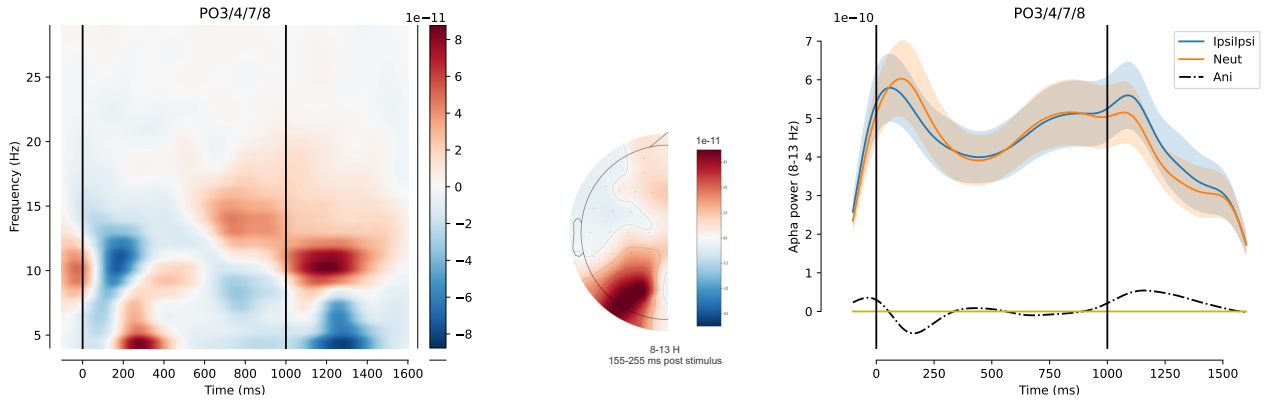
with time measurement 190-235, so it does not need to report the result of that. Moreover, even though, in the regression model, the main effect of Anc was highly significant (t=3.377, p<0.01) and the main effect of N2nc were not significant (t=1.549, p=0.150), the Pearson correlation coefficient disclosed the relationship between RT and N2nc (r=0.780, p<0.001). For non-decision time parameter of contralateral versus neutral in model$_t$ as the best model, difference non-decision time were predicted by N2nc and Anc (R-square=0.428,F(2,13)=4.114, p <0.05). In addition, the main effect of N2nc was significant (t=2.746, p<0.05), and the main effect of Anc was not significant (t=-0.504, p=0.624). For accuracy of contralateral versus neutral, the difference of accuracy was not predicted by N2nc and Anc (R-square=0.123,F(2,13)=0.7701, p=0.486). Therefore, for ipsilateral versus neutral, there were no predictions, see Table 2. Therefore, the visualization of the result are reported by Component-Component plus Residual (CCPR) just for the non-decision time parameter and RT regression models with N2nc and Anc, and non-decision time regression model with N2ni and Ani, see Figure 7.

3 Discussion

This work investigated how spatial prioritization leads to response faster than non-spatial prioritization in the perceptual decision making task. In order to explain the elaborate framework in the behavioral, cognitive process, and electrophysiological levels for top-down intention, the contribution post-stimulus informative and uninformative N2 subcomponent and alpha power have been considered. Therefore, contralateral minus uninformative N2 subcomponent and alpha power predicted the behavioral and the relevant DDM parameters. Furthermore, we hypothesized that: 1. spatial attention may not shift any other resources during the top-down signal. 2. It might be connected to the information accumulation process. 3. It should be related to the speed of decoding or motor response (non-decision time). 4. It may lead to bias to the locus of the presented stimulus. 5. It could shift the amount of information that is needed to decide. Generally, all nested hypotheses fed to the robust and rigorous model comparisons to choose the best relevant parameter. Then, model number three (model$_t$) won based on two powerful measures DIC and R-square. In addition, lateralized central sites relating to motor response and parietal-occipital relating to decoding stimulus, so these sites were candidates to control the top-down signal. Therefore, we investigated the contribution of alpha power lateralization and N2 sub-component amplitudes to spatial attention in perceptual decision making. As a result, the current study revealed that two regressors could predict the non-decision parameter of the model$_t$ and RT associated with the spatial cue, see Tables 1 and 2. The informative (contralateral or ipsilateral) signal allocate some resources to speed up the procedures of perceptual decision making. Nonetheless, an uninformative signal has a neutral effect and does not manipulate other resources to progress the speed of cognitive processing. Therefore, the main question is which informative lateralizations are highly important to flourish behavioral performance during prioritization in perceptual decision making. Consequently, what we unveiled only the difference between contralateral and neutral components which could predict segregatedly the difference non-decision time ($T_{er,P-NP}$ = T_{er}(P) - T_{er}(NP)) and RT (RT_{P-NP} = RT(P) - RT(NP)) between prioritization and non-prioritization. This finding confirms that the mechanism of contralateral oscillation to the attended spatial attention is more



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



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

Figure 6. Time–frequency lateralization. (a) The grand-average of contralateral minus neutral single-trial time–frequency wavelets are indicated at electrodes PO3/4/7/8 on the left hand. 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. Therefore, the ±50 msec around of 50% FAL on the time window from -100 msec to 600 msec of the stimulus onset has been shown by the red color 237-337 msec after the stimulus onset. The topographic map is based on differences of 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 illustrated at the posterior-occipital electrodes on the left hand. Moreover, the mean of contralateral and neutral alpha power as well as the outcome of (ipsilateral minus neutral) are displayed on the right hand. Therefore, the result of FAL has been shown by the red color 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. Moreover, the first and second vertical lines show cue prioritization and stimulus presence respectively, and shaded regions identify the standard error.

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]	t=-2.746 p<0.05	2.048e+07(1.32e+07) [8.62e+06 4.96e+07]	t= 1.549 p=0.150	-2.999e+06(2.42e+06) [-8.32e+06 2.32e+06]	t=-1.241 p=0.241
Anc	-3.485e+07(6.92e+07) [-1.87e+08 1.17e+08]	t=-0.504 p=0.624	1.8e+11**(5.33e+10) [6.27e+10 2.97e+11]	t=3.377 p<0.01	5.19e+09(9.74e+09) [-1.63e+10 2.66e+10]	t=0.533 p=0.605
R-square	0.428		0.679		0.123	
F-statistic	F(2,13) = 4.114, p < 0.05		F(2,13) = 11.66, p < 0.01		F(2,13) = 0.7701, p =0.486	

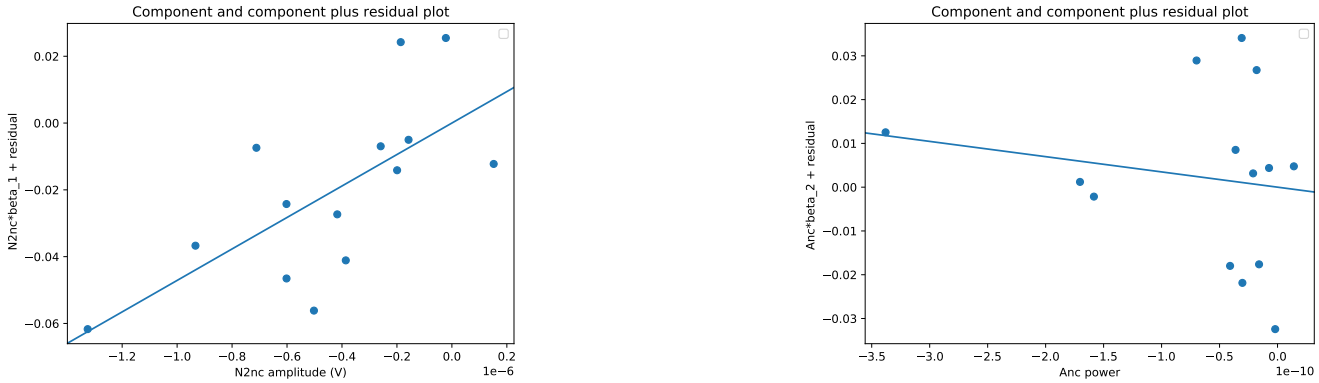
T_{P-NP} = T_{er}(yes) - T_{er}(no), RT_{P-NP} = RT(P) - RT(NP), Acc_{P-NP} = accuracy(P) - accuracy(NP), SE = standard error, CI = confidence interval. R^2 and beta-regression are given for separate linear regression model, yes= prioritization, no= no prioritization. Asterisks indicate significant p value as *p<0.05, **p<0.01.

Table 1. Separate multiple linear regression between the dependent variable (non-decision time of prioritization minus non-prioritization, RT and accuracy) and predictors (N2nc and Anc), FDR correction for two different N2nc, t-test statistics for each Beta parameters, confident interval, standard error.

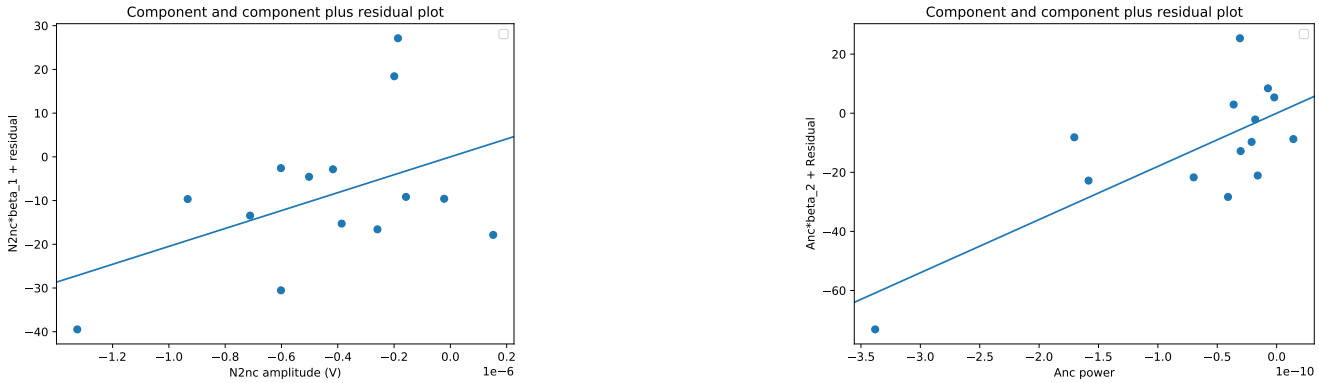
Output	$T_{er,P-NP}$		RT_{P-NP}		Acc_{p-NP}	
Predictors	β (SE) [95% CI]	t	β (SE) [95% CI]	t	β (SE) [95% CI]	t
Intercept	-0.0194(0.013) [-0.047 0.008]	t=-1.548 p=0.150	-11.6572(10.140) [-33.976 10.662]	t=1.1700(1.326) p=0.275	t=0.882 [-1.749 4.089]	p=0.397
N2ni	8286.7188(2.52e+04) [-4.71e+04 6.36e+04]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

T_{P-NP} = T_{er}(P) - T_{er}(NP), RT_{P-NP} = RT(P) - RT(NP), Acc_{P-NP} = accuracy(P) - accuracy(NP), SE = standard error, CI = confidence interval. R^2 and beta-regression are given for separate linear regression model, yes= prioritization, no= no prioritization. Asterisks indicate significant p value as *p<0.05, **p<0.01.

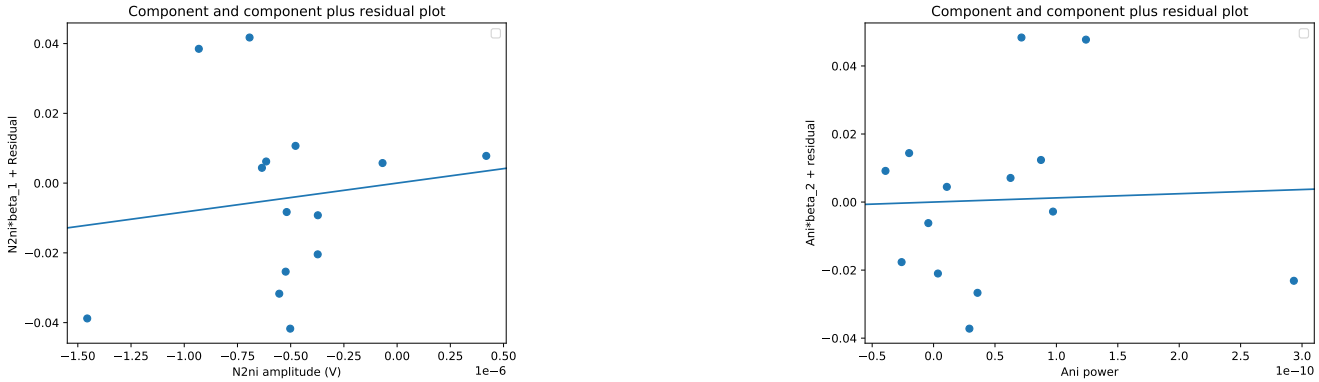
Table 2. Separate multiple linear regression between the dependent variable (non-decision time of prioritization minus non-prioritization, RT and accuracy) and predictors (N2ni and Ani), FDR correction for two different N2ni, t-test statistics for each Beta parameters, confident interval, standard error.



(a) Non-decision time regression model with N2nc and Anc.



(b) Reaction time regression model with N2nc and Anc.



(c) Non-decision time regression model with N2ni and Ani.

Figure 7. Component-Component plus Residual plots, a. For the non-decision parameter model, the main effect of the N2nc amplitude is revealed (t=2.746, p<0.05) and Anc was no significant (t=-0.504, p=0.624), b. For the RT time regression model, the main effect of the Anc power is found (t=3.377, p<0.01) and the N2nc was not significant (t=1.549, p=0.150). But, the difference RT was also highly correlated with N2nc by Pearson correlation coefficient (r=0.780, p<0.001) which had not seen in the regression model, c. The N2i (t=0.329, p=0.748) and Ani (t=0.100, p=0.922) were not significant effects in non-decision time. Moreover, all regressors were not significant effects in accuracy mode in both RT and non-decision time regression.

significant than ipsilateral oscillation. Most previous works ([7,24,25,59]) have concentrated on the difference of contralateral and ipsilateral deflection and have not explored which relationship informative and uninformative could predict relevant the DDM parameters associated with spatial attention. Innovatively, in the current study, we also found a significant relationship between non-decision time and contralateral minus neutral activity ([7,22,60]).

3.0.1 N2nc and Anc band predict reaction time

Behavioral data revealed a significant effect of spatial prioritization relating to non-prioritization. Therefore, in order to decipher the mechanism of the difference, the multiple regression model using N2nc (the time measurement 260-315 msec) and Anc band as regressors could highly predict the difference RT between prioritization and non-prioritization, see Table 1. The effect of N2nc and Anc coefficients were displayed in Figure 7b. Also, the components of N2ni and Ani together could not explain the variance of the difference RT, see Table 2. Consequently, this result shows that contralateral activity could interpret the different spatial cues in comparison with ipsilateral activity. According to post-stimulus of figure 5, both N2 subcomponent contralateral and ipsilateral ERP were more negative than neutral ERP in line with latter findings ([15,17]). Nevertheless, post-stimulus contralateral alpha power was more negative than neutral alpha power, but ipsilateral alpha power was more positive than neutral alpha power, see 6. Negative Anc shows that alpha contralateral oscillation to the attend visual field was increased as an excitatory process because our brain made resources the most sensitive to decode stimulus from the attended area. On other hand, positive Ani shows that alpha ipsilateral deflection to the attend visual field was increased as an inhibitory process because our brain endeavored to control and attenuate the non-attended spatial location ([72]). In fact, our brain applied these two contradictory mechanisms to facilitate the speed of perceptual decision making. Moreover, the regression and statistical analyses proved that the role Anc and N2nc were more significant to predict the different RT of spatial attention. So that, the more N2nc and Anc led to the more difference RT associated with spatial attention. Now, the next question research is how the cognitive process and the DDM parameter might interpret this significant effect, and whether the DDM model fitting could provide further elucidation or not.

3.0.2 N2nc amplitude predicts one of the cognitive modeling parameters

Now, the lack of information about the DDM parameters and neural correlates in spatial orienting of perceptual decision making satisfied us to elucidate deeply about it. We also explored the measure of alpha lateralization and N2nc in visual perceptual decision making. As the aforementioned, the difference in RT of prioritization and non-prioritization originated from the difference of non-decision time parameter based on the nested model comparisons. Interestingly, we could show that N2nc and Anc predicted the non-decision parameter based on the nested model comparison, see Table 1. In addition, the main effect of N2nc was significant, but Anc was not significant, see Figure 7a. Oppositely, N2ni and Ani could not predict non-decision time parameter, see Table 2 and Figure 7c. As a result, the N2nc and as Anc might reflect the non-decision parameter and RT. So that, the role of N2nc was more worthwhile in order to interpret the non-decision parameter difference. According to our literature review, this finding is the first research to show the role of N2nc in the variance within a spatial cue, its behavioral performance, and non-decision parameter. It also shows that the contralateral activity was superior to elucidate the effect of the spatial cue in comparison with ipsilateral activity. Therefore, the more N2nc and Anc caused the more difference of non-decision time relating to spatial attention. Finally, the remaining question is if contralateral minus neutral alpha power has had the main role to reflect the variance RT within two different spatial cues, what is its role in predicting the DDM parameters? In this study, we applied the strategy of post-stimulus components such as alpha lateralization and N2 subcomponent to predict the variance RT and the DDM parameter and did not check other strategies such as pre-stimulus, then it might be the answer.

4 Conclusion

The non-decision time parameter could reflect the spatial cue in perceptual decision making based on the DDM model fitting. Present findings propose that the amplitude of N2nc at the central site predicted RT as well as the non-decision time of spatial prioritization. Also, the Anc power oscillations at posterior-occipital only predicted RT of spatial prioritization and it unrelated to non-decision time parameter. In addition, none of the N2ni amplitude and Ani power could not predict RT and non-decision time between spatial prioritization and non-prioritization. Therefore, the role of contralateral activity was worthwhile to allocate the brain's resources in order to fascinate and accelerate the decision making process in comparison with ipsilateral activity. The more N2nc led to the more difference of non-decision time associated with spatial attention, but the more N2nc or Anc led to the more difference of RT associated with spatial attention. In addition, the non-decision time parameter might interpret perfectly the modulation of the attentional orienting during decision making process.

Acknowledgement

All data used in this research is publicly available in the open science framework (https://osf.io/q4t8k/). The authors are very grateful to Dirk Ostwald and his colleagues that make this data available. The hDDM implementation code and simulation

scripts for all current cognitive models are readily available at https://github.com/AGhaderi/hDDM_attention.

References

1. Gold, J. I. & Shadlen, M. N. The neural basis of decision-making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
2. Gherman, S. & Philiastides, M. G. Human vmpfc encodes early signatures of confidence in perceptual decisions. *eLife* **7**, 1–18 (2018).
3. Shadlen, M. N. & Kiani, R. Decision making as a window on cognition. *Nature* **80**, 791–806 (2013).
4. Philiastides, M. G., Ratcliff, R. & Sajda, P. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* **26**, 8965–8975 (2006).
5. Ostwald, D., Porcaro, C., Mayhew, S. D. & Bagshaw, A. P. EEG-fMRI based information theoretic characterization of the human perceptual decision system. *PLoS ONE* **7**, e33896 (2012).
6. Yeshurun, Y. & Carrasco, M. Spatial attention improves performance in spatial resolution tasks. *Vis. Res.* **39**, 293–306 (1999).
7. Klatt, L. I. *et al.* Unraveling the relation between eeg correlates of attentional orienting and sound localization performance: a diffusion model approach. *J. Cogn. Neurosci.* **32**, 945–962 (2020).
8. Krajbich, I., Lu, D., Camerer, C. & Rangel, A. The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* **3**, 193 (2012).
9. Nunez, M. D., Vandekerckhove, J. & Srinivasan, R. How attention influences perceptual decision making: Single-trial eeg correlates of drift-diffusion model parameters. *J. mathematical psychology* **76**, 117–130 (2017).
10. S. Gluth and M. S. Spektor and J. Rieskamp. Value-based attentional capture affects multi-alternative decision making. *eLife* **7**, e39659 (2018).
11. Posner, M. I. Orienting of attention. *The Q. journal experimental psychology* **32**, 3–25 (1980).
12. Sagar, V., Sengupta, R., Devarajan & Sridharan. Dissociable sensitivity and bias mechanisms mediate behavioral effects of exogenous attention. *Sci. Rep.* **9**, 1–13 (2019).
13. Imani, E., Harati, A., Pourreza, H. & Goudarzi, M. M. Brain-behavior relationships in the perceptual decision-making process through cognitive processing stages. *Neuropsychologia* **155**, 107821 (2021).
14. Praamstra, P. & Oostenveld, R. Attention and movement-related motor cortex activation: a high-density eeg study of spatial stimulus–response compatibility. *Cogn. Brain Res.* **16**, 309–322 (2003).
15. Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K. & Awh, E. Alpha-band oscillations enable spatially and temporally resolved tracking of covert spatial attention. *Psychol. Sci.* **28**, 929–941 (2017).
16. Li, Y., Lou, B., Gao, X. & Sajda, P. Post-stimulus endogenous and exogenous oscillations are differentially modulated by task difficulty. *Front. Hum. Neurosci.* **7**, 1–10 (2013).
17. Rihs, T. A., Michel, C. M. & Thut, G. Mechanisms of selective inhibition in visual spatial attention are indexed by α-band eeg synchronization. *Eur. J. Neurosci.* **25**, 603–610 (2007).
18. van Schouwenburg, M. R., Zanto, T. P. & Gazzaley, A. Spatial attention and the effects of frontoparietal alpha band stimulation. *Front. human neuroscience* **10**, 658 (2017).
19. Praamstra, P., Boutsen, L. & Humphreys, G. W. Frontoparietal control of spatial attention and motor intention in human eeg. *J. neurophysiology* **94**, 764–774 (2005).
20. Haegens, S., Luther, L. & Jensen, O. Somatosensory anticipatory alpha activity increases to suppress distracting input. *J. cognitive neuroscience* **24**, 677–685 (2012).
21. Bernier, P. M., Whittingstall, K. & Grafton, S. T. Differential recruitment of parietal cortex during spatial and non-spatial reach planning. *Front. human neuroscience* **11**, 249 (2017).
22. Praamstra, P. Prior information of stimulus location: effects on erp measures of visual selection and response selection. *Brain research* **1072**, 153–160 (2006).
23. Amenedo, E., Lorenzo-Lopez, L. & Pazo-alvarez, P. Response processing during visual search in normal aging: the need for more time to prevent cross talk between spatial attention and manual response selection. *Biol. psychology* **91**, 201–211 (2012).

24. Gamble, M. L. & Luck, S. J. N2ac: An erp component associated with the focusing of attention within an auditory scene. *Psychophysiology* **48**, 1057–1068 (2011).

25. Cespon, J., Galdo-Alvarez, S. & Diaz, F. Cognitive control activity is modulated by the magnitude of interference and pre-activation of monitoring mechanisms. *Sci. reports* **6**, 1–11 (2016).

26. Ratcliff, R. & McKoon, G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation* **20**, 873–922 (2004).

27. Olianezhad, F., Zabbah, S., Tohidi-Moghaddam, M. & Ebrahimpour, R. Residual information of previous decision affects evidence accumulation in current decision. *Front. Behav. Neurosci.* **13**, 1–12 (2019).

28. Birte U. Forstmann and Roger Ratcliff and Eric-Jan Wagenmakers. Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annu. review psychology* **67**, 641–666 (2016).

29. Smith, P. L. Diffusion theory of decision making in continuous report. *Psychol. Rev.* **123**, 425–451 (2016).

30. Ratcliff, R. A theory of memory retrieval. *Psychol. review* **85**, 59 (1978).

31. Stone, M. Models for choice-reaction time. *Psychometrika* **25**, 251–260 (1960).

32. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion decision model: Current issues and history. *Trends Cogn. Sci.* **20**, 260–281 (2016).

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

34. Evans, Nathan J and Hawkins, Guy E and Brown, Scott D. The role of passing time in decision-making. *J. experimental psychology: learning, memory, cognition* **46**, 316 (2020).

35. Evans, N. J. & Brown, S. D. People adopt optimal policies in simple decision-making, after practice and guidance. *Psychon. Bull. & Rev.* **24**, 597–606 (2017).

36. Evans, N. J., Bennett, A. J. & Brown, S. D. Optimal or not; depends on the task. *Psychon. bulletin & review* **26**, 1027–1034 (2018).

37. Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N. & Pouget, A. The cost of accumulating evidence in perceptual decision making. *The J. Neurosci.* **32**, 3612–3628 (2012).

38. Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D. & Wagenmakers, E.-J. Bayesian parametric estimation of stop-signal reaction time distributions. *J. Exp. Psychol. Gen.* **142**, 1047 (2013).

39. Gomez, P., Ratcliff, R. & Perea, M. A model of the go/no-go task. *J. Exp. Psychol. Gen.* **136**, 389 (2007).

40. Usher, M. & McClelland, J. L. Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* **111**, 757–769 (2004).

41. P. D. Kvam. A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *J. Math. Psychol.* **91**, 14–37 (2019).

42. Mallahi-Karai, K. & Diederich, A. Decision with multiple alternatives: Geometric models in higher dimensions - the cube model. *J. Math. Psychol.* **93**, 102294 (2019).

43. Pedersen, M. L., Frank, M. J. & Biele, G. The drift diffusion model as the choice rule in reinforcement learning. *Psychon. bulletin & review* **24**, 1234–1251 (2017).

44. Fontanesi, L., Gluth, S., Spektor, M. S. & Rieskamp, J. A reinforcement learning diffusion decision model for value-based decisions. *Psychon. bulletin & review* **26**, 1099–1121 (2019).

45. D. K. Sewell and H. K. Jach and R. J. Boag and C. A. Van Heer. Combining error-driven models of associative learning with evidence accumulation models of decision-making. *Psychon. bulletin & review* **26**, 868–893 (2019).

46. Gluth, S., Kern, N., Kortmann, M. & Vitali, C. L. Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nat. Hum. Behav.* **4**, 634–645 (2020).

47. Ratcliff, R. Decision making on spatially continuous scales. *Psychol. Rev.* **125**, 888–935 (2018).

48. Choi, W. & Paik, S.-B. Intrinsic timescales of sensory integration for motion perception. *Sci. Reports* **9**, 1–15 (2019).

49. Wagenmakers, E.-J., Van Der Maas, H. L. & Grasman, R. P. An ez-diffusion model for response time and accuracy. *Psychon. bulletin & review* **14**, 3–22 (2007).

50. Brown, S. & Heathcote, A. A ballistic model of choice response time. *Psychol. review* **112**, 117 (2005).

51. Usher, M. & McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. review* **108**, 550 (2001).

52. Tillman, G., Van Zandtc, T. & Loganb, G. D. Sequential sampling models without random between-trial variability: the racing diffusion model of speeded decision making. *Psychon. bulletin & review* (2020).

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

54. Voss, A., Nagler, M. & Lerche, V. Diffusion models in experimental psychology: A practical introduction. *Exp. psychology* **60**, 385 (2013).

55. R. Ratcliff and J. N. Rouder. Modeling response times for two-choice decisions. *Psychol. Sci.* **9**, 347–356 (1998).

56. Voss, A., Rothermund, K. & Voss, J. Interpreting the parameters of the diffusion model: An empirical validation (2004).

57. Ratcliff, R. & Tuerlinckx, F. Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychon. bulletin & review* **9**, 438–481 (2002).

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

59. Luck, S. J. An introduction to the event-related potential technique. *J. Chem. Inf. Model.* **53** (2014).

60. Gramfort, A. *et al.* Meg and eeg data analysis with mne-python. *Front. Neurosci.* **7** (2013).

61. Cohen, M. X. A better way to define and describe morlet wavelets for time-frequency analysis. *NeuroImage* **199**, 81–86 (2019).

62. Vandekerckhove, J., Tuerlinckx, F. & Lee, M. D. Hierarchical diffusion models for two-choice response times. *Psychol. Methods* **16**, 44–62 (2011).

63. Wiecki, T. V., Sofer, I. & Frank, M. J. Hddm: Hierarchi cal bayesian estimation of the drift-diffusion model in python. *Front. neuroinformatics* **7**, 14 (2013).

64. Ratcliff & Childers, R. Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision* **2**, 237 (2015).

65. Spilcke-Liss, J., Zhu, J., Gluth, S., Spezio, M. & Glascher, J. Semantic incongruency interferes with endogenous attention in cross-modal integration of semantically congruent objects. *Front. Integr. Neurosci.* **13**, 53 (2019).

66. Ratcliff, R., Philiastides, M. G. & Sajda, P. Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the eeg. *Proc. Natl. Acad. Sci.* **106**, 6539–6544 (2009).

67. Philiastides, M. G., Heekeren, H. R. & Sajda, P. Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. *J. Neurosci.* **34**, 16877–16889 (2014).

68. Gamerman, D. & Lopes, H. F. (eds.) *Markov chain monte carlo: stochastic simulation for bayesian inference* (Taylor and Francis, London, 2006).

69. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**, 457–472 (1992).

70. Brooks, S. P. & Gelman, A. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* **7**, 434–455 (1998).

71. Spiegelhalter, D. J., Carlin, B. P. & Linde, A. V. D. Bayesian measures of model complexity and fit. *J. Royal Stat. Soc. Ser. B (Statistical Methodol.* **64**, 583–639 (2002).

72. Akiko Ikkai and Sangita Dandekar and Clayton E. Curtis. Lateralization in alpha-band oscillations predicts the locus and spatial distribution of attention. *PLoS One* **11**, e0154796 (2016).

Supplementary Material

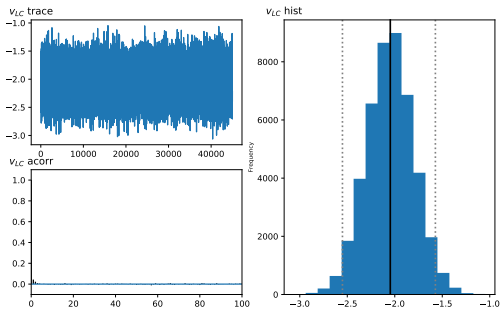
Model fitting and Model comparison

In this paper, five possible models were run to give us appropriate intuition of spatial prioritization. To compare which model is better to fit observed data, our analysis was equipped with two powerful weapons called DIC score (lower is better) and R-square (near to one is better). It should be noted that a difference in DIC scores of 15 and above is considered meaningful[65,71]. Therefore, the result of model comparison analysis reveals that model$_t$ provided the best model fit so that it received the smallest DIC and largest R-square, so non-decision time parameter (T_{er}) interpret perfectly the manipulation of spatial prioritization, see Table 3.

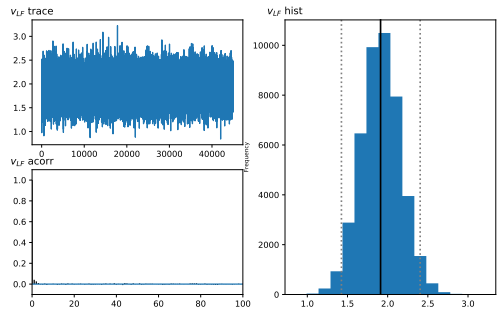
To assess model convergence (i.e. model$_t$), the MCMC sampling traces of the group posteriors distribution for these parameters which show converged chain procedure during parameters estimation were utilized, see Figures 8, 9 and 10. There were horizontal lines (convergence) in the trace and no jump in the auto-correlation plots, and the group mean posteriors were smooth with no jagged signs.

Model	$v \sim$ coher + stim	$v \sim$ spat	$T_{er} \sim$ spat	$z \sim$ spat	$a \sim$ spat	DIC	R^2(RT)	R^2(Acc)
$model_p$	$\surd$					-3536.59	0.902	0.716
$model_v$	$\surd$	$\surd$				-3555.87	0.915	0.768
model$_t$	$\surd$		$\surd$			**-3610.76**	**0.940**	**0.888**
$model_z$	$\surd$			$\surd$		-3525.91	0.770	0.371
$model_a$	$\surd$				$\surd$	-3562.36	0.773	0.579

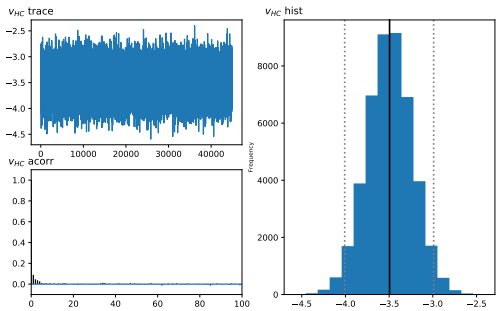
Table 3. Model comparison based on DIC for all model variants tested in this study and R2 metrics for RT and accuracy in the group level. coher, coherency; stim, stimulus; spat, spatial attention; RT, reaction time; Acc, accuracy.



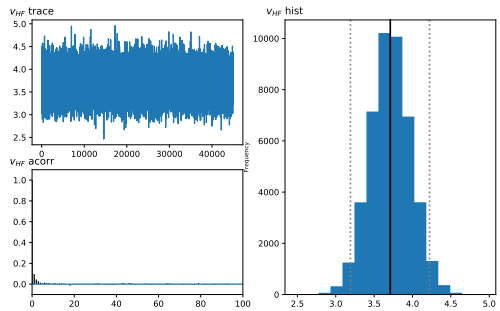
(a) Low coherency low and car stimulus.



(b) Low coherency low and face stimulus.

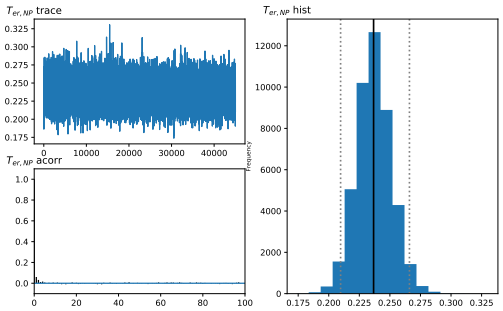


(c) High coherency high and car stimulus.

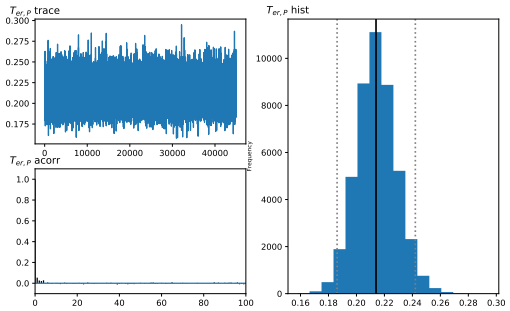


(d) High coherency and face stimulus.

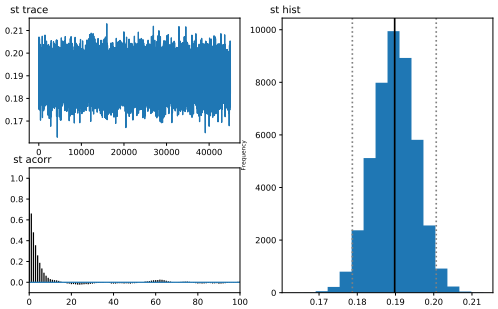
Figure 8. Plots of model convergence evaluation with the trace, auto-correlation, and histogram posterior distribution of drift rate parameter for the low coherency (upper half) and high coherency (bottom half) conditions of both types of car and face stimuli.



(a) Non-decision time parameter T_{er} for non spatial prioritization.

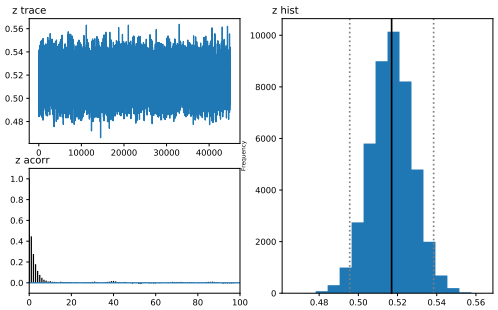


(b) Non-decision time parameter T_{er} for spatial prioritization.

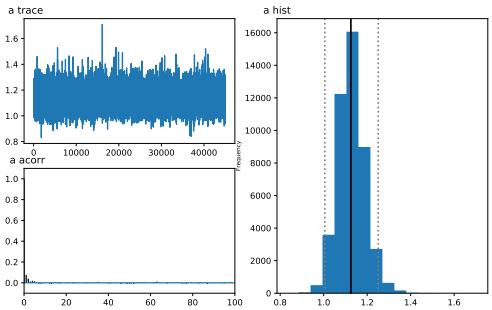


(c) Non-decision time variability parameter s_t.

Figure 9. Plots of model convergence evaluation with the trace, auto-correlation, and histogram posterior distribution of non-decision time parameter for the prioritization (top left) and no prioritization coherency (top right) conditions. Non-decision time variability is also reported in the bottom half.



(a) Starting point parameter z.



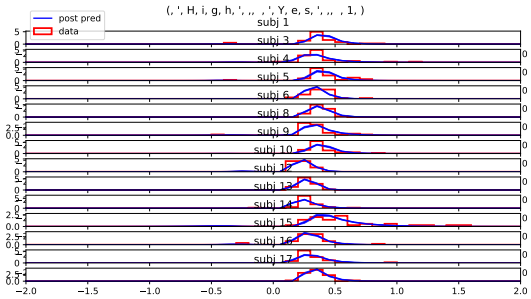
(b) Boundary decision parameter a.

Figure 10. Plots of model convergence evaluation with the trace, auto-correlation, and histogram posterior distribution of (a) starting point parameter and (b) boundary decision parameter.

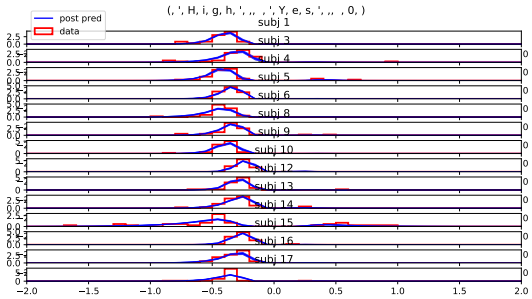
R-hat values for all models: The convergence of the Markov chains was assessed by calculating the R-hat Gelman-Rubin statistic ([69]) to ensure that the models had properly converged, which compares between-chain and within-chain variance. The model of "model$_t$" showed superior convergence and the same stationary distribution with four chains, based on R-hat values under 1.0001 as follow:

model$_t$:
'$T_{er,NP}$':0.9999935100922528,
'$T_{er,P}$': 0.9999916748628642,
'v_{HC}': 1.0000426841659067,
'v_{HF}': 0.999991945644348,
'v_{LC}': 1.0000080181533444,
'v_{LF}': 1.0000028712137732,
'a': 0.9999968891198591,
'st': 1.0000144779358509,
'z': 1.0000630870848732,

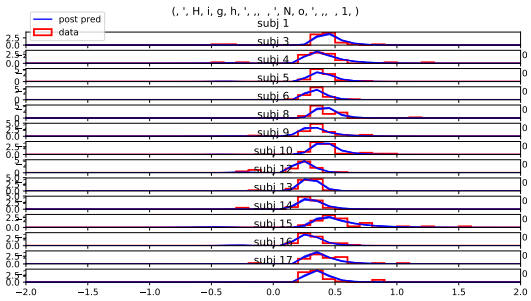
Plot of posterior predictive against observed data: To assess the model fit, we compared posterior model predictions with the observed data. In fact, posterior predictive checks of the best model were conducted by averaging 500 simulations generated from the model's posterior to confirm it could reliably reproduce patterns in the observed data. There was a good agreement between the observed data and the model predictions across conditions in all conditions. Overall, the model is able to capture the overall shape of reaction times across the four different conditions, see Figure 11.



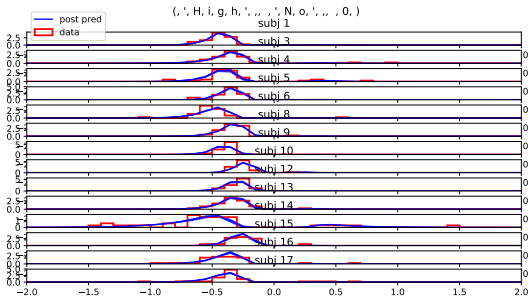
(a) High coherence, prioritized and face stimulus.



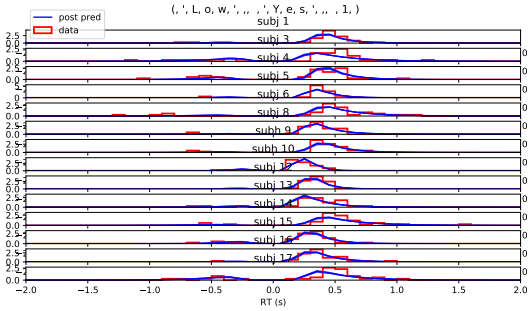
(b) High coherence, prioritized and car stimulus.



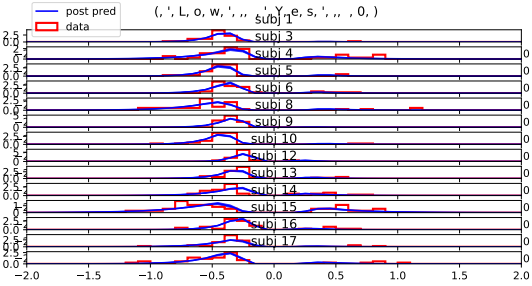
(c) High coherence, non-prioritized and face stimulus.



(d) High coherence, non-prioritized and car stimulus.

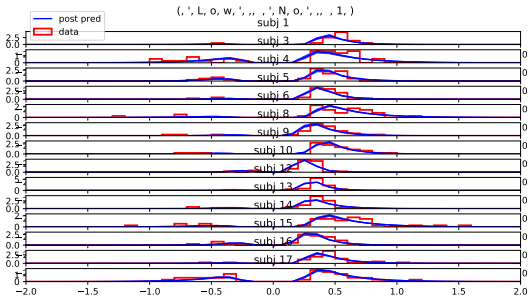


(e) Low coherence, prioritized and face stimulus.

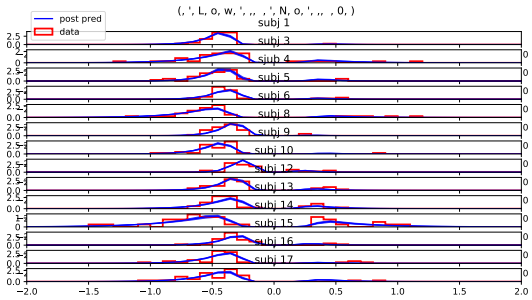


(f) Low coherence, prioritized and car stimulus.

h



(g) Low coherence, non-prioritized and face stimulus.



(h) Low coherence, non-prioritized and car stimulus.

Figure 11. The histogram shows observed reaction time distribution for all participants across all conditions (positive vs. negative results on the x-axis correspond to face vs. car stimuli). The line plots the results from the posterior predictive simulation.