CORTEX 190 (2025) 262—285

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Effects of alertness on perceptual detection and discrimination

Yanzhi Xu [a,*], *Martijn Wokke* [a,b], *Valdas Noreika* [c,a], *Corinne Bareham* [a], *Sridhar Jagannathan* [a], *Stanimira Georgieva* [a], *Caterina Trentin* [a] and *Tristan Bekinschtein* [a]

[a] *Consciousness and Cognition Laboratory, Department of Psychology, University of Cambridge, UK*
[b] *Department of Experimental Psychology, Mind, Brain, and Behavior Research Center (CIMCYC), University of Granada, Spain*
[c] *Department of Biological and Experimental Psychology, School of Biological and Behavioural Sciences, Queen Mary University of London, UK*

ARTICLE INFO

Article history:
Received 11 March 2025
Revised 5 June 2025
Accepted 30 June 2025
Action editor Céline R. Gillebert
Published online 10 July 2025

Keywords:
Perceptual decision-making
Alertness
Psychometric curve
Signal detection theory

ABSTRACT

The level of alertness fluctuates throughout the day, exerting modulatory effects on human cognitive processes at any moment. However, our knowledge of how alertness level interacts with specific cognitive demands and perceptual rules of a task is still limited. Here we used perceptual decision-making paradigms to explore this issue. We analysed data from four different experiments involving a total of 113 participants: 1) auditory masking detection, 2) sensorimotor detection, 3) auditory spatial discrimination, and 4) auditory phoneme discrimination. We examined participant performance during the natural transition from awake (high alertness) to drowsy (low alertness). First, we fitted psychometric functions to the performance in EEG-defined high and low alertness metastable states. Second, we modelled slope and threshold from the fitted sigmoidal curves as well as signal detection theory measures, including perceptual sensitivity (d') and response bias (criterion). We found lower detection and discrimination sensitivity to stimuli as alertness level decreases, signalled by a shallower slope and a lower d', while the threshold increases slightly and equivalently across experiments. We observed no change in criterion during the transition. Zooming in, we observed that the decrease in sensitivity measured by slope was stronger for discrimination than for detection decisions, indicating that lower alertness impairs the precision of decisions in discriminating alternatives more than in identifying the presence of a stimulus around the threshold. Taken together, these results suggest that alertness has a common effect on perceptual decision-making and differentially modulates detection and discrimination decisions.

---

* *Corresponding author*. Consciousness and Cognition Lab, Department of Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, UK.
E-mail address: yx333@cam.ac.uk (Y. Xu).

https://doi.org/10.1016/j.cortex.2025.06.018

1. Introduction

Perceptual decision-making refers to the process of detecting, discriminating, and categorising information from the senses (Hanks & Summerfield, 2017). It is a fundamental cognitive function that underlies a wide range of everyday behaviours, from checking if a charger's indicator light is on to judging whether a traffic light is yellow or red. Because of its clear behavioural manifestations and its central role in cognition, perceptual decision-making is often regarded as a window to human cognition (Shadlen & Kiani, 2013). However, it remains unclear how this capability fluctuates with alertness—a general state of readiness that enables adaptive responses to environmental events (Moller et al., 2006; Posner, 2008). Alertness has received increasing attention due to its pervasive influence on performance. Research on this interplay is important not only for understanding human consciousness and cognition but also for improving psychological and physiological experiments by controlling for fluctuating levels of alertness, which often introduces noise in measurements. Furthermore, this interaction has real-world implications for health, well-being, and society, particularly for shift workers and long-distance drivers, who often make decisions while in a state of reduced alertness. Despite its importance, the complex interactions between alertness levels and cognitive processes require systematic investigation (Goupil & Bekinschtein, 2012; Mediano et al., 2021). In this study, we address this gap by examining how fluctuations in alertness influence perceptual decision-making across four independent experiments involving detection and discrimination tasks.

1.1. *Perceptual decision-making*

Perceptual decision-making is a core component of human cognition, involving weighing sensory information and making judgments based on internal decision rules. This process is fundamental to human interaction with the environment, enabling the transformation of sensory inputs into actionable responses. It also plays a central role in shaping subjective experience by determining how ambiguous sensory input is interpreted. Importantly, because perceptual decisions are often simplified into binary judgments on sensory perception, they offer a foundational model for understanding general decision mechanisms. The core perceptual and evaluative mechanisms revealed by perceptual decision-making can help elucidate the broader principles underlying more complex cognitive processes (O'Connell et al., 2018).

Two commonly used types of experiments in the study of perceptual decision-making are detection and discrimination tasks. In a detection task, a target stimulus is presented with a certain probability and the subjects are required to detect the presence of the target. Subjects are asked to respond 'yes' or 'no' based on the information they perceive. The hit rate, the percentage of 'yes' responses when the stimulus is presented, is usually used to measure performance. In comparison, there are two stimuli in a discrimination task. Subjects are asked to make judgments between two stimuli based on the signal strength they perceive for each alternative. The accuracy, the percentage of correctly choosing one stimulus when it is presented, is the primary behavioural measure of discrimination.

The hit rate and accuracy are general behavioural measures, which are the manifestation of a series of inner processes, including sensory processing, stimuli recognition, decision-making processing and execution of motor commands. The discrepancy in processing surely exists between detection and discrimination tasks, however, these two types of tasks share the common elements of decision formation that we are specifically interested in this paper. One can consider detection as a comparison between a stimulus and noise while discrimination as comparing between two stimuli. The common decision-making stage is the accumulation of sensory evidence until reaching some threshold level (or criterion) and triggering the commitment to a decision (Gold & Shadlen, 2007; Samaha et al., 2020). Empirically, research has shown that contrast thresholds in yes/no detection and orthogonal discrimination tasks are indistinguishable and the two tasks are functionally equivalent (Smith & Ratcliff, 2009; Thomas & Gille, 1979).

To explore perceptual decision formation, models in perceptual psychophysics have been deployed, including psychometric functions and signal detection theory (SDT). A commonly used psychometric function is the sigmoidal function which depicts the relationship between signal intensity and performance (Del Cul et al., 2007; Noreika et al., 2020a; Sandberg et al., 2011). This relation is further characterised by threshold and slope, two parameters in the sigmoidal function model. Slope of the sigmoidal curve reflects the precision of decision (i.e., how sharply performance changes with stimulus intensity), while threshold indicates the stimulus level required for a given performance level, often interpreted as perceptual bias (Prins, 2016). Previous research has shown that decreased alertness leads to shallower slope and increased psychophysical threshold (Noreika et al., 2017).

However, the psychometric approach has limitations, as it confounds perceptual and decisional factors, relying on a single performance measure (hit rate). In contrast, SDT better captures the distinction between the perception of sensory evidence and the criterion for making a judgment (Green & Swets, 1966; Macmillan & Creelman, 2005). By analysing hit and false alarm rates, it dissociates two independent components from performance: perceptual sensitivity (d') and response criterion (C). Reduced alertness has been found to correlate with impaired perceptual sensitivity while leaving the criterion unaffected (Jagannathan et al., 2022). Crucially, SDT's dissociation between sensitivity and bias provides a more comprehensive account of perceptual decision-making than psychometric slope/threshold parameters alone.

1.2. *Fluctuation in alertness*

As we fall asleep each night, we gradually lose consciousness of the world and ourselves. Low alertness lies between full wakefulness and early sleep, unfolding as a dynamical continuum of neurobehavioural changes. Crucially, we distinguish alertness from related constructs: consciousness (in the sense of awareness of both the external world and internal states), arousal (the physiological activation dependent on circadian factors and sleep pressure; Borbély et al., 2016), and vigilance (sustained forced-monitoring attention). Low alertness marks the transition from being aware and awake—characterised by wakefulness and responsiveness (Bekinschtein et al., 2009)—to a state of reduced responsiveness and a lower sensitivity to external stimulation (Goupil & Bekinschtein, 2012). However, these changes in alertness do not equate to alterations in consciousness or arousal alone. For example, anesthesia reduces arousal and eliminates consciousness completely, whereas sleep onset may retain some awareness despite low arousal. Similarly, Steghaus and Poth (2024) showed that subjective states of relaxation and sleepiness, though both low-arousal states, are qualitatively distinct. Moreover, alertness is distinct from vigilance: while alertness reflects a state of physiological readiness, vigilance refers to task-specific fluctuations in attentional engagement, which can decline due to factors unrelated to sleepiness such as boredom (Hancock, 2013).

Alertness fluctuations can be categorised into two types. Tonic alertness represents slow, intrinsic fluctuations of alertness over minutes to hours, driven by factors like fatigue, circadian rhythms, or task engagement; while phasic alertness refers to a rapid readiness state, triggered by external cues like warnings (Poth, 2025). While tonic and phasic alertness appear to operate on distinct timescales, they interact dynamically: phasic alertness enhances performance most effectively at intermediate tonic levels as described by the Yerkes-Dodson law (an inverse U-shaped function; Yerkes & Dodson, 1908). This interaction is crucial when investigating the effects of alertness on cognition. Alertness in terms of behaviour or experience is measured after stimulus onset, and is thus likely contaminated by external stimuli or motor responses. In contrast, pre-trial neural markers can serve as cleaner indicators of baseline alertness state, as they precede task-evoked processes and are less influenced by phasic fluctuations.

Low tonic alertness occurs not only before nighttime sleep onset but also frequently during the daytime (Carrier & Monk, 2000; Goel et al., 2011). Most previous research on alertness has focused on pupil-linked arousal (McGinley et al., 2015; van Kempen et al., 2019; Waschke et al., 2019), which primarily reflects the early stages of transition from wakefulness to drowsiness. Few studies have examined fluctuations across the full spectrum of alertness levels and their relationship to cognitive processes (Goupil & Bekinschtein, 2012), despite the pervasiveness of alertness fluctuations throughout the day. Here we explore a broader range of alertness levels, from full wakefulness to eyes-closed sleep onset, to more comprehensively examine how alertness affects cognitive functions.

The transition from high to low alertness does not occur in a moment but is a continuum of changes. The complex process of falling asleep involves the convergence of physiological, EEG and behavioural dynamics (Ogilvie, 2001). In addition to physiological changes like reduced heart rate and respiratory activity, one outstanding feature of lower alertness is the change in the frequency spectrum of EEG signals. According to the Hori scoring system that assesses the depth of drowsiness (Hori et al., 1994), the electrical brain activity is dominated by the alpha (α) band during relaxed wakefulness. As people get into mild drowsiness, the alpha wave trains become less frequent and the EEG activity may even become flat. Instead, another oscillation with a higher amplitude, theta (θ) wave, emerges. Finally, when people enter deeper stages in the transition, vertex sharp waves, spindles and K-complexes appear successively. These EEG changes are closely linked to cognitive performance and subjective experience: alpha dropout and theta emergence or an increased θ/α rate correlate with slower reaction times (Noreika et al., 2020a), more omissions (Jagannathan et al., 2018), higher error rates (Bareham et al., 2014; Jung et al., 1997), as well as self-reported sleepiness (Diaz et al., 2016). In this study, we employ two electrophysiological methods to classify alertness levels based on the above EEG characteristics of drowsiness: micro-measures and the θ/α power ratio (see details below).

At the behavioural level, a more obvious characteristic of sleep onset is that subjects gradually lose the ability to respond to external stimuli, including the diminished performance of accuracy, reaction time and response rate (Bareham et al., 2014; Noreika et al., 2020a; Ogilvie, 2001). Researchers have used short reaction times to indicate wakefulness and longer reaction times to mark decreasing alertness (Ogilvie et al., 1991). An important caveat is the potential confounding influence of vigilance. While shorter reaction times are commonly associated with increased alertness (Canales-Johnson et al., 2020), they can also reflect reduced vigilance over time-on-task (Esterman et al., 2013; Fortenbaugh et al., 2017). On the other hand, extremely slow reactions often represent lapses in sustained attention (Unsworth & Robison, 2016). As individuals move deeper into metastable states during the transition, their responsiveness diminishes or completely ceases (Harsh et al., 1994). This loss of responsiveness is often considered evidence of progression into the later stages of N1 and the onset of true N2 sleep (Ogilvie, 2001). These behavioural dynamics are also correlated with EEG changes during the transition to sleep. Reaction time increased along Hori stages, while response cessation is usually found during later Hori stages (Hori et al., 1994; Liberson & Liberson, 1966). Therefore, we complement EEG markers with behavioural indicators of reaction times and the proportion of omissions, to ensure a valid operationalisation of alertness.

1.3. *Current study*

Previous research has examined how alertness affects perceptual decision-making by manipulating signal intensity. Lower signal intensity (representing externally driven noise) typically reduces the probability of stimulus awareness in detection tasks or impairs stimulus identification in discrimination tasks (Koch & Preuschoff, 2007; Marcel, 1983; Sergent & Dehaene, 2004). This reflects how external noise (lower intensity) impairs our cognitive performance. In contrast, decreased alertness introduces internal noise, which interacts with external stimuli to influence cognitive performance. Noreika et al. (2017, 2020a) demonstrated how alertness levels modulate perceptual decision-making across varying signal intensities in detection tasks, while Jagannathan et al. (2022) showed that decreased alertness biases spatial discrimination by shifting the subjective midline. Although detection and discrimination share a two-alternative structure, they may engage distinct underlying cognitive mechanisms (Bridwell et al., 2013; Correa et al., 2004; Smith & Ratcliff, 2009). However, it is still unknown whether the alertness fluctuation exerts a common effect on detection and discrimination decision-making or if it interacts with the specific cognitive demands and perceptual rules of the task at hand.

In this study, we are interested in the effect of alertness level on decision-making performance in simple detection and discrimination tasks. In other words, how the detection and discrimination performance changes as people go from fully awake to low alertness metastable states. To this end, we conduct original secondary analyses of existing experimental data from four studies, three of which have been published. These studies include two detection tasks: (1) auditory masking (Noreika et al., 2020a) and (2) sensorimotor response (Noreika et al., 2017; 2020b), as well as two discrimination tasks: (3) spatial attention (Jagannathan et al., 2022) and (4) phoneme discrimination. Building upon these established paradigms, we use both psychometric function (focusing on slope and threshold) and SDT (analysing sensitivity d' and criterion C) to investigate the change in the properties of decision-making in detection and discrimination tasks during the natural transition from wakefulness to low alertness.

This study employs the following methods. First, all trials in each experiment are divided into high alertness or low alertness based on 4 different methods: a) an automatic micro-measures algorithm (Jagannathan et al., 2018) that uses EEG data of 4 sec before the stimulus onset to capture the variance and coherence features of the EEG frequency space and recognise the patterns of elements like vertex, K complex, and spindles; b) θ/α ratio on the 2-s pretrial signal; c) 45 percent median split on fast and slow reaction times (RTs); d) proportion of missed responses within 10 consecutive trials. Second, changes in performance indices between alertness levels are investigated. At each alertness level, the hit rate of target stimuli with varied intensities is modelled per participant using a classic psychometric sigmoidal function. Then slope and threshold obtained from the sigmoidal function for each participant are compared between high and low alertness within each experiment. Further, we apply SDT to analytically decouple perceptual sensitivity from response biases, to characterise the mechanisms underlying the behavioural changes. Last but not least, we use multilevel modelling (MLM) to test whether different perceptual demands have an impact on the relationship between alertness and decision-making performance.

The hypotheses tested in this study stem from the preregistration (https://osf.io/3p4rf). We hypothesise that alertness level has a common effect on detection/discrimination for all four experimental settings: as people enter lower alertness, their precision of decision (slope of the sigmoidal curve) decreases, while their threshold (the stimulus intensity at the inflexion point of the sigmoidal curve) may increase. The hypotheses from the preregistration are as follows:

H1. We hypothesise that the proportion of correct responses to detect or discriminate between stimuli increases with stimulus intensity following a sigmoidal curve, regardless of conscious state.

H2. As people become drowsy, their general sensitivity (the slope of the sigmoidal curve) decreases [as seen in Noreika et al. (2020b)], which may be attributed to reduced perceptual sensitivity (d'). Specifically:

a) As alertness decreases, detection sensitivity decreases.
b) As alertness decreases, discrimination between stimuli sensitivity decreases.

H3. As people become drowsy, their threshold (the stimulus intensity at the inflection point) may increase. This is determined by an increase in the criterion relative to the sensory distribution.

a) Drowsy states will increase the detection threshold.
b) Drowsy states will increase the discrimination threshold.

H4. Common effect of alertness: Alertness level has a main effect on detection/discrimination for all four experimental settings, both in threshold (increase) and slope (decrease).

H5. Alertness parametrically affects decision-making. The drowsier the participant, the greater the performance parameter (e.g., proportion correct, threshold, slope, and criterion) change.

H6. (exploratory): Alertness may change the psychophysics curve. Exploratorily, we expect the goodness of fit to a sigmoidal curve to decrease with lower alertness as compared to wake.

H7. (exploratory): Alertness slows the reaction of the participants (Bareham et al., 2014, 2015; Canales-Johnson et al., 2020; Comsa et al., 2019; Goupil & Bekinschtein, 2012; Noreika et al., 2020b). We will explore associations between reaction times and changes in performance measures for the combined data of all four experiments.

We mainly address the first four hypotheses in this study. Hypothesis 5 was intended to explore the parametric effects of alertness; however, due to a substantial imbalance in the distribution of drowsiness levels—particularly the low number of 'late drowsy' trials—we were unable to reliably test this hypothesis as we treated alertness as a binary variable to ensure sufficient statistical power (see details in the main analysis). Exploratory hypotheses 6 and 7 are discussed in the supplementary materials, in section 2 and 13, respectively.

2. Methods

2.1. *Participants and experiments*

We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. Sample sizes and the inclusion/exclusion criteria were determined by the previous research prior to data analysis. Further exclusions after alertness classification were described in the 'Measures of alertness' section.

Four studies were analysed to explore the interaction between the level of alertness and perceptual decision-making performance (Fig. 1). The first two studies were detection tasks and the other two were discrimination tasks. These are explained in detail as follows but in short, they included auditory detection of a target masked by white noise, kinaesthetic detection of a muscle movement induced by transcranial magnetic stimulation (TMS), auditory spatial discrimination, and phoneme morphing discrimination. For all experiments, participants were seated comfortably with their eyes closed in a dimmed room and were encouraged to relax during these tasks, allowing the low alertness state to emerge while responding to the stimuli or cues. For the two discrimination tasks, in addition to the above drowsy session, there was an awake session when participants were seated upright with lights on and instructed to stay awake. High and low alertness were classified within the drowsy session for detection tasks and were determined between awake and drowsy sessions in discrimination tasks. For both high and low alertness levels, we assess participants' performance according to varied stimuli intensities. The difference in performance among experiments is further examined in order to test a common effect of alertness level on cognitive processing.

2.1.1. *Experiment 1 (auditory masking detection)*

This data was collected to investigate the influence of the wakefulness state on auditory detection, and the basic threshold and slope but not SDT results were already published (Noreika et al., 2020a). 56 adults were recruited to participate in an auditory detection task through the electronic volunteer database of the MRC Cognition and Brain Sciences Unit at the University of Cambridge. They were screened using the Epworth Sleepiness Scale (Johns, 1991) with a minimum sleepiness score of 7 in order to be likely to become drowsy in the experiment. Following screening and subsequent exclusions due to poor performance (e.g., high false alarm rates in catch trials or poor response fitting), 31 participants remained in the study (9 male; mean age 27.4; age range 20–39). During the 2-h experiment, participants relaxed and were presented with a

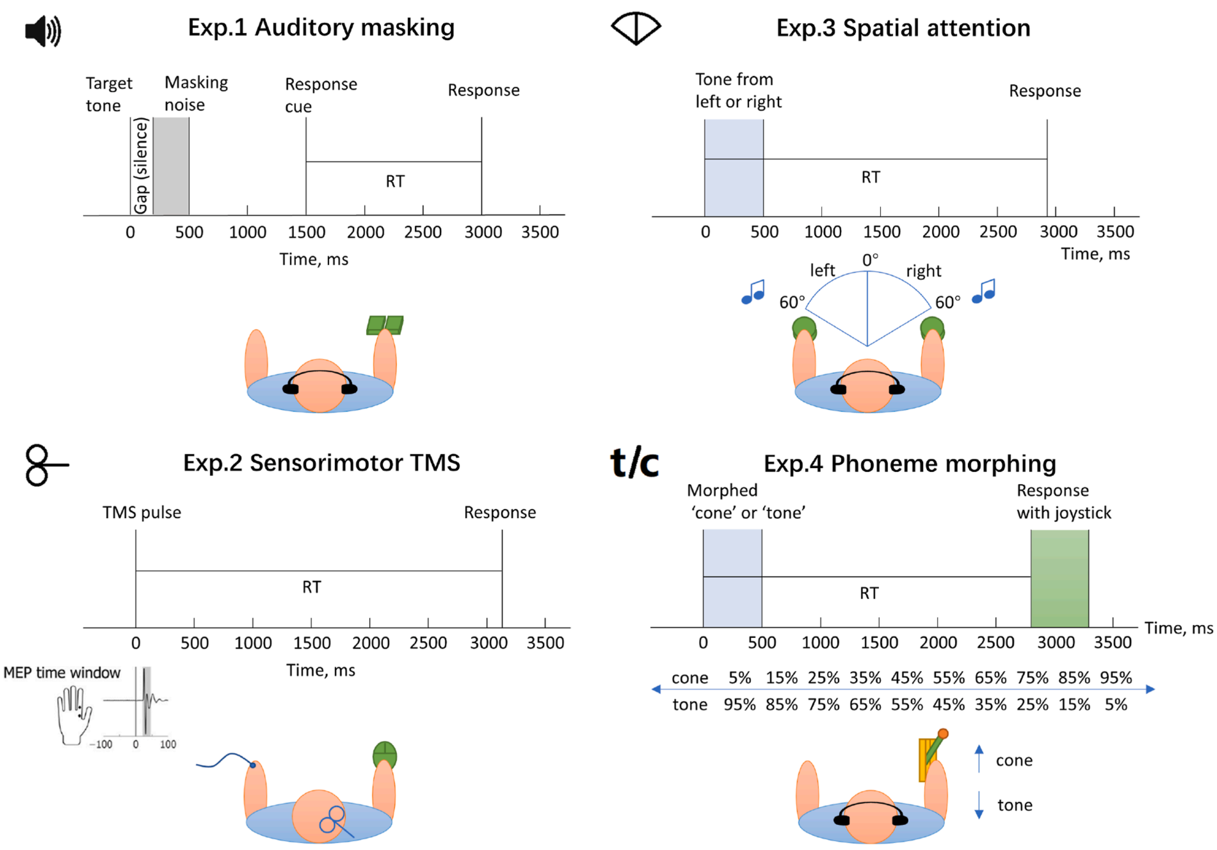

Fig. 1 – Experimental design. The experimental design of four tasks is shown (the left two are detection tasks and the right two are discrimination tasks). The first task used a target tone masked by a noise with a varying gap. Participants were asked to respond whether they heard the target tone after a response cue. The second experiment was a sensorimotor task, different from the other three tasks in using auditory modality. TMS with different intensities around individual thresholds was applied to the right motor cortex of participants. They responded using their right hand to indicate whether they felt a twitch in the left hand. In the third task, participants were presented with the target tones that fell on their left or right side and responded to indicate the location of the target with left or right buttons. The last task used the morphed phoneme of 'cone' and 'tone' as the target and participants were asked to move a joystick to indicate the word they heard. The distance they moved the joystick represented their confidence in the answer. By including these tasks, we covered a wide range of perceptual decision-making processes in the testing of the effects of alertness levels.

target sound (10 msec; 1000 Hz; 2.5 msec fade-in and fade-out; attenuation: −24 Db) and a following masking sound of white noise (300 msec; 707.1–1414.2 Hz; 5 msec fade-in and fade-out; attenuation: 0 Db), between which the duration was varied near the individual threshold (Fig. 1: Exp. 1). The individual threshold (the duration between target sound and masking noise) was estimated before the experiment. Participants were asked to report if they heard a target sound by hitting one of two buttons after a response cue. No response within 6 sec from the onset of the target stimulus was defined as unresponsive. There were 11 conditions of the interval between the target stimulus and the noise (0, .25, .5, .75, .875, 1, 1.125, 1.25, 1.5, 1.75, 2 * individual threshold). Additional catch trials were also performed where there was no target sound before the mask. On average, 501 trials were carried out per participant (*SD* = 65, *Min* = 361, *Max* = 604). EEG signals were recorded during the whole experiment. The experimental protocol was approved by the Cambridge Psychology Research Ethics Committee, and participants received a remuneration of £30 for the study.

2.1.2. *Experiment 2 (sensorimotor detection to TMS)*

The experiment was conducted to explore how cortical evoked responses to TMS are modulated by alertness level, with basic psychometric results (fitted using a logistic function) reported previously (Noreika et al., 2017; 2020b). Here, we re-analysed the same data using a different psychometric function (described below) and additionally applied SDT, which was not included in the original studies. 20 right-handed participants (7 male; mean age 23.7; age range 21–33) attended this experiment. They were also screened with the Epworth Sleepiness Scale (Johns, 1991) and the mean score was 9.4 (*SD* = 4.3). 17 participants were included in the analysis after excluding those with too low responding rates. TMS was applied to the right motor cortex with the output intensity varying around the individual resting motor threshold (−20%, −15%, −10%, −5%, 0%, +5%, +10%, +15%, +20%), which was determined using a criterion of $\geq 50\ \mu V$ motor-evoked potential (MEP) amplitude in at least five out of ten consecutive trials (Ikoma et al., 1996; Rossini et al., 1994; Samii et al., 1996). Participants were instructed to respond whether they felt a twitch or a touch in their left hand after each TMS pulse by clicking mouse buttons with their right hand (Fig. 1: Exp. 2). No response within 6 sec after TMS was considered an omission. 520 trials were performed for each participant with inter-pulse intervals of 8.5–10.5 sec. In addition to TMS-evoked potentials (TEP) at the site of stimulation and MEP from the first dorsal interosseous (FDI) muscle of the left hand, scalp EEG activities and Surface Electromyography (EMG) from the FDI of both hands were also recorded during the experiment. TEP, MEP and MEG are not analysed in this paper because we focus on behavioural performance here. The experimental protocol was approved by the Medical Research Ethics Committee of The University of Queensland and each participant received a remuneration of $30.

2.1.3. *Experiment 3 (spatial attention discrimination)*

In this experiment, complex harmonic tones were used to investigate the effect of alertness level on spatial attention. Psychometric and SDT results of the experiment were published (Jagannathan et al., 2022). 41 right-handers were recruited for this study. After discarding invalid data due to technical problems or not following instructions, 32 participants (14 male; 24.46 ± 3.72 years old) remained in the analysis, among which 29 of them had a sleepiness score ≥7 (easy sleepers) and 3 had a sleepiness score ≥4 of the Epworth Sleepiness scale (Johns, 1991). All participants attended two sessions (awake and drowsy sessions). The awake session lasted about 8 min and consisted of 124 trials. The drowsy session lasted 1.5–2 h and included 740 trials. Both sessions contained the same auditory tone localisation task in which participants were presented with complex harmonic tones from left 59.31° to right 59.31° of their spatial midline to guide their spatial attention (Fig. 1: Exp. 3). 12 steps of 4.01° were used between 11.17° and 59.31° in both left and right directions, and smaller steps of 1.86° were used between left 11.17° and right 11.17°. Participants were instructed to report the direction of the tone (left/right) with a button press. The response window was 5 sec. EEG was also recorded during the whole experiment. The study was approved by the Cambridge Psychology Research Ethics Committee and participants received £30 for taking part in the study.

2.1.4. *Experiment 4 (phoneme morphing discrimination)*

33 right-handed participants (from an initial pool of 45) were included in the analysis after excluding those with insufficient data in the low alertness state. They were presented with a mixture of morphed spoken words ‘tone’ and ‘cone’. Ten different percentages of morphing between the two words were used as conditions of difficulty (Fig. 1: Exp. 4). The participants reported whether they had heard the word ‘tone’ or ‘cone’ and their confidence in the answer by moving a joystick. The direction of the movement (forwards or backwards) indicated the word choice and the extent to which they moved indicated confidence. According to our hypotheses, here we analyse the choice of participants but not their confidence. No feedback was provided throughout the task. The experiment consisted of a 12-min awake session (160 trials) and a roughly 85-min drowsy session (480 trials). During the awake session, the response window was 2.5 sec and inter-stimulus intervals varied between 2.5 and 4 sec. In the drowsy session, the response window was 4 sec and inter-stimulus intervals ranged between 4 and 8 sec. Participants were woken up when they had missed 3 consecutive responses. The study was approved by the ethical committee of the Medical Research Council for the Cambridge Brain Sciences Unit.

2.2. *EEG acquisition and preprocessing*

2.2.1. *Experiment 1 (auditory masking)*

128-channel EEG data were recorded at a sampling rate of 500 Hz using the Net Amps 300 amplifier (Electrical Geodesics Inc., Oregon, USA). After excluding channels over forehead, cheeks, and neck, 92 channels were retained in the analysis. Raw data were filtered between .5 Hz and 40 Hz, re-referenced to the average, and epoched between −4000 msec and 6000 msec relative to the onset of the target sound. The preprocessing was conducted in MATLAB (The MathWorks) using the EEGLAB toolbox (Delorme & Makeig, 2004).

2.2.2. Experiment 2 (sensorimotor TMS)

For this TMS-EEG combined experiment, 64-channel EEG data were recorded using the BrainAmp MR Plus amplifier, TMS BrainCap, and Brain Vision Recorder (v1) software (Brain Products; Gilching, Germany). The recording was downsampled to 500 Hz and epoched between −4000 msec and −10 msec relative to each TMS pulse for calculating EEG spectral power before TMS.

2.2.3. Experiment 3 (spatial attention)

During both awake and drowsy sessions, 128-channel EEG data were recorded at a sampling rate of 500 Hz (Electrical Geodesics Inc., Oregon, USA). Channels over forehead, cheeks and neck were excluded, retaining 92 channels in the analysis. Raw data were filtered between 1 Hz and 40 Hz and epoched between −200 msec and 800 msec relative to the onset of the tone stimuli as post-trial epochs. Additionally, data were epoched between −4000 msec and 0 msec to the onset of the stimuli in the drowsy session for classifying the alertness level.

2.2.4. Experiment 4 (phoneme morphing)

EEG data were recorded at a sampling rate of 1000 Hz by using 128-channel HydroCel Sensors and a GES300 amplifier (Electrical Geodesics Inc., Oregon, USA), and 92 channels were retained in the analysis. Raw data were filtered between .5 Hz and 40 Hz, re-referenced to the average of all electrodes, and epoched between −200 msec and 2000 msec relative to the onset of the target sound. Drowsy session data were also epoched between −4000 msec and 0 msec to the onset of the stimuli for classification of alertness levels.

Furthermore, for all experiments, independent component analysis (ICA) was performed to remove artefacts and noisy channels were interpolated.

2.3. Measures of alertness

The first step in the analysis is to classify the alertness level of subjects during the experiments. There is no single widely accepted measure of lower alertness in cognitive experiments. Therefore, we used a combination of four methods to validate and assess the consistency between methods employed to define lower alertness: 1) micro-measures based automatic classification of alertness levels based on multiple parameters including EEG power/coherence and spatio-temporal signatures (Jagannathan et al., 2018); 2) relative change in θ/α spectral power ratio; 3) relative change in reaction time length; 4) proportion of missed trials. In detection tasks, trials were divided into 'high alertness' and 'low alertness' using the above methods. In discrimination experiments, since they contained extra awake sessions, we classified all trials in awake sessions as 'high alertness' and lower alertness trials in drowsy sessions as 'low alertness'. We did not apply alertness classification methods to the awake sessions because these sessions were conducted under a controlled experimental setting specifically designed to promote wakefulness, including: explicit instructions to stay awake, short inter-stimulus intervals, reduced overall session duration, constant lighting, and upright seating. This definition of 'high alertness' represents a truly higher alertness state compared to high alertness trials in drowsy sessions. Still, we performed complementary analyses using only drowsy sessions in discrimination tasks to match with detection tasks and confirm the results (Supplementary section 4, 7, 10, and 12). Participants with too few data in any alertness state after classification were excluded from further analysis.

2.3.1. Micro-measures

The micro-measures algorithm is an automated method to detect micro variations in levels of alertness in EEG experiments under eye-closed settings (Jagannathan et al., 2018). It applies a support vector machine (SVM) and individual element detectors on the 4-s pretrial EEG data to classify each trial into different alertness levels. Specifically, two major steps were conducted to generate the micro-measures algorithm. First, predictor variance and coherence features were computed and then used to classify data into 'awake' and 'drowsy' using the SVM. Individual alpha band range (individual alpha peak ± 2 Hz) was used in the computation. Second, the drowsy data were further examined using individual element detectors to detect vertex, K-complex, and spindles. Then another SVM was used to detect true spindles based on the variance and coherence features. Data with vertex, K-complex, and true spindles were identified as 'late drowsy', while other drowsy data were marked as 'early drowsy'.

In our data, the average number of 'late drowsy' trials per subject was 24, which was far fewer than the 'early drowsy' trials (288 on average). Therefore, without further analysing multiple levels of alertness (Hypothesis 5), we combined both drowsy conditions into a single low alertness category and then compared it with high alertness state (awake). Following classification, several participants were excluded from subsequent analyses due to insufficient numbers of trials in either high or low alertness state. As a result, the final number of participants included in experiments 1, 2, and 3 was 23, 16, and 31, respectively, while no participants were excluded in Experiment 4.

Since micro-measures states were computed before stimulus onset, they were not confounded by cognitive processes related to external stimuli or motor responses. This method has been validated to outperform both manual scoring of the Hori scale and the θ/α ratio (Jagannathan et al., 2018). Therefore, we used the micro-measures algorithm as the primary method to classify alertness levels, and included the θ/α ratio along with two behavioural markers (described below) as complementary measures.

2.3.2. EEG θ/α power

Previous research has shown that lower alertness can be characterised by increased theta band power and decreased alpha band power (Hori et al., 1994), or an increased ratio of θ/α (Bareham et al., 2014; Noreika et al., 2020a). Here, we computed a ratio of θ/α power based on the pretrial data, according to which we split the data into equal proportions of high alertness and lower alertness trials for each participant. More specifically, data from 2-s pretrial epochs with respect to the onset of stimuli were used to compute the spectral power of EEG frequency oscillations. θ and α power were extracted respectively from the results of complex Morlet wavelet convolution (Cohen, 2014), in which nine frequencies logarithmically increased between 4 Hz and 12 Hz and the number of wavelet cycles was 10. Convolution results were first averaged over −2 sec–0 sec. Then data of 4–7 Hz were averaged to obtain θ power and data of 8–12 Hz were averaged to calculate α power for each electrode and each trial. The θ/α ratio was then averaged across all electrodes. In this way, we obtained a single θ/α value for each trial. Data for each participant were divided into equal proportions of lower 45% and higher 45% θ/α power ratio as high alertness and low alertness, excluding 10% intermediate trials.

2.3.3. *Reaction times (RTs)*

The alertness level can also be measured by RT and RT variability (Bareham et al., 2014; Ogilvie, 2001). Studies show that people respond slower as they become drowsy compared to awake (Hori et al., 1994; Ogilvie et al., 1989). To apply the RTs measure, the same criterion of equal proportions was used as θ/α power ratio. Trials with 45% shortest RTs were labelled as high alertness and trials with 45% longest RTs were marked as low alertness for each participant. The advantage of the RTs measure is that it reveals the alertness level at the exact time of response rather than seconds before the response as the micromeasures and θ/α ratio measure. However, some research shows no association between RT and lower alertness (Baulk et al., 2001). So we cautiously use the RTs measure as a complementary tool. To reduce potential confounding influences such as lapses and impulsive responses, we limited trial durations, after which responses were considered as omissions, and excluded extremely short reaction times from analysis.

2.3.4. *The proportion of omissions*

As alertness decreases, responses of subjects become progressively intermittent (Lagarde & Batejat, 1994; Makeig et al., 2000; Ogilvie et al., 1989). The lack of response to the stimuli could also be used as a measure of sleepiness (Williams et al., 1959). Previous research on lower alertness inspected the sequence of omissions and marked omission as a lack of response within a 6-sec window after stimulus presentation (Comsa et al., 2019). In this paper, we inspected the proportion of missed responses to indicate subjects' alertness level. Behavioural data were divided into blocks of 10 trials and the number of unresponsiveness within each block was counted. Blocks with 2 or more omissions were marked as low alertness and blocks with no omissions were marked as high alertness (Bareham et al., 2014). Blocks with 1 omission were considered intermediate trials and were excluded from further analysis. This method of alertness measurement is therefore even more diluted in time. After classification, participants with too few trials in either alertness state were excluded from further analyses, yielding final sample sizes of 20, 16, 24, and 25 for experiments 1 through 4, respectively.

2.4. *Signal detection theory analysis*

SDT has been widely used in the analysis of detection and discrimination experiments. The significant contribution of SDT is that it disentangles the encoding process and decision-making process by computing two parametric statistics: the perceptual sensitivity (d’) and criterion (C) (Green & Swets, 1966). The sensitivity d’ quantifies the ability to discriminate signals from noise while C captures response bias independent of sensitivity (Macmillan & Creelman, 2005). SDT models perceptual decisions as a process where observers compare a decision variable representing the strength of sensory evidence against a criterion. When the decision variable exceeds this criterion, a positive response is generated. This decision therefore serves as a stage that connects the representation of stimuli to behavioural responses. Here, d’ and C were calculated based on behavioural performance [Equations (1) and (2)] for each participant and each level of alertness and further modelled by alertness and experiment later.

$$d' = z(H) - z(F) \tag{1}$$

$$C = -\frac{1}{2}[z(H) + z(F)] \tag{2}$$

where H represents the hit rate, F represents the false alarm rate, and z stands for transformation to a z score.

More specifically, for the auditory masking task, the false alarm rate was calculated based on catch trials and the hit rate was obtained from the pooled performance across different conditions of difficulty. For the TMS task, the weakest TMS pulse was used to calculate the false alarm rate since there were no catch trials in this experiment. To make the parameters comparable among detection and discrimination tasks, we treated one stimulus in the discrimination tasks as the target and another stimulus as noise. Therefore, the weakest condition for the target stimulus was used to derive the false alarm and all other conditions were pooled together to get the hit rate in the two discrimination experiments. For the spatial attention task, the ‘left’ responses were considered as hits, based on the evidence of reduced performance for left-sided stimuli during decreased alertness (Bareham et al., 2014; Jagannathan et al., 2022)—a phenomenon attributed to the selective impairment of spatial processing in drowsiness (Bareham et al., 2015). Defining hits this way helps to avoid the left bias when calculating the false alarm rates in SDT analyses. For the phoneme morphing discrimination task, we randomly used ‘cone’ as the target stimulus and used 5% ‘cone’ condition to obtain the false alarm rate[1].

2.5. *Psychometric curve fitting*

To describe the change in detection and discrimination with different alertness levels, we fitted the psychometric curves to the hit rates of target stimuli with different intensities in high and low alertness trials, respectively. Since the birth of SDT (Green & Swets, 1966) and its use in experimental psychology, several methods have been developed to fit the performance

---

[1] In a subset of the discrimination data, the target and noise stimuli were shifted and tested, showing no difference from the current results for all four behavioural measures in the phoneme morphing task and for three measures in the spatial attention task except for criterion. The criterion in the spatial attention task decreased with lower alertness when using right-sided stimuli as the target and it was most likely due to the left bias.

curves. One way is to model psychometric functions with SDT (Prins, 2016). Since d’ and the proportion of correct responses could be mutually converted, we could construct a function that describes the relationship between stimulus intensity and d’, and then further convert it to behavioural performance. Another way of fitting a psychometric curve is to use sigmoidal functions that describe the relationship between stimuli and performance directly, with parameters of slope and threshold (Prins, 2016). Sigmoidal functions commonly used include cumulative normal distributions, logistic functions, and Weibull functions. The Weibull function is given as Equation (3):

$$F(x;\alpha,\beta) = 1 - exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) \tag{3}$$

where α is the threshold corresponding to $F(x = \alpha;\ \alpha, \beta) = 1 - exp(-1^{\beta}) \approx .632$ and β is slope. We used the Weibull function because the former two are inappropriate when stimulus intensity $x = 0$ represents an absence of signal (Prins, 2016). The psychometric function is further formulated as Equation (4):

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)F(x;\alpha,\beta) \tag{4}$$

where γ corresponds to the guess rate (lower asymptote), and λ corresponds to the lapse rate (upper asymptote).

We fitted the performance curves using the SDT method and Weibull function with the Palamedes toolbox for Matlab (Prins & Kingdom, 2018). The goodness of fit for both methods was 1 in a test conducted on the first experiment using micro-measures split. However, the Weibull function better captured the sigmoidal shape of the actual performance across varying stimulus intensities in our data. Therefore, we opted for Weibull fitting in the subsequent analysis. One possible explanation for the difference is that the number of trials may not have been sufficient to provide a reliable estimation of d’ in each condition of difficulty for SDT psychometric fitting. Besides the choice of the fitting function, the setting for the guess rate γ and lapse rate λ can also influence the fitting outcomes (Green & Swets, 1966; Wichmann & Hill, 2001). The guess rate refers to the probability of a positive response when stimuli are undetected, while the lapse rate corresponds to the percentage of errors for highly detectable stimuli (Gold & Ding, 2013). To minimise the influence of these rates on our target parameters, slope and threshold, we set the guess and lapse rates to be the same for high and low alertness levels. We first combined the performance under high and low alertness to fit a curve and determine common guess and lapse rates. These common values were then fixed for separate fittings of performance in high and low alertness (except for the spatial attention task with a left bias, we set lapse rates independently for high and low alertness to better fit the data). This approach ensured that different alertness levels shared the same guess and lapse rate settings, allowing us to focus clearly on comparing slope and threshold. If effects were found under these constrained settings, we would expect even greater effects under unconstrained conditions. Additionally, we fitted the individual psychometric curves using three different settings: individual best parameters, group best parameters, and fixed parameters (with guess and lapse rates set to .05). In the main text, we present results using individual best rates, while results from group best rates (group fitting) and .05 settings (strict fitting) are included in the supplementary material. Importantly, the different asymptote values did not affect the overall result patterns, strengthening the confidence in our findings on alertness effects on slope and threshold.

2.6. *Multilevel modelling*

To investigate the common effect of alertness on detection and discrimination decision-making, MLM was used on the above behavioural measures (slope, threshold, d’ and criterion). The difference between high and low alertness was first illustrated using raincloud plots (Allen et al., 2021) with colours from cbrewer (Harrower & Brewer, 2003), and statistically compared within each experiment. Next, MLM was performed on the aggregated data of four experiments using the lmer function from the lme4 package (Bates et al., 2015) in R (R Core Team, 2013), to compare the difference in the behavioural performance between high and low alertness across different experimental settings. Subject identity was used as a random factor, and alertness level and experiment were used as fixed factors. Models with different combinations of fixed factors (alertness alone, experiment alone and both alertness and experiment) were compared using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the negative log-likelihood. The F test for the complete model (alertness * experiment) was further conducted by using the anova function.

This study’s hypotheses were preregistered; see https://osf.io/3p4rf. An earlier version of the manuscript was published as a preprint on bioRxiv (Xu et al., 2023). The data and analysis code used in the paper to create evidence are available at https://doi.org/10.17863/CAM.95292.

3. Results

In the following sections, we present statistical and modelling results on four behaviour measures: slope and threshold from psychometric curve fitting, and d’ and criterion from SDT. We concentrate on the effects of alertness captured by the micro-measures method and compare them, in the text, to the other three alertness split methods (see the supplementary material for further results and specific figures). Additional analyses on slope and threshold using different curve fitting parameters, as well as slope analyses using integrated sessions in discrimination tasks, are also discussed, and detailed results are shown in the supplementary material.

3.1. *The precision of decision decreases in low alertness states*

For both states of alertness, performance follows a sigmoidal curve as predicted in hypothesis H1. Critically, with decreased alertness, the curves are markedly shallower (lower slope) as compared to the high alertness across all experiments (see Fig. 2). This suggests that participants are less sensitive to

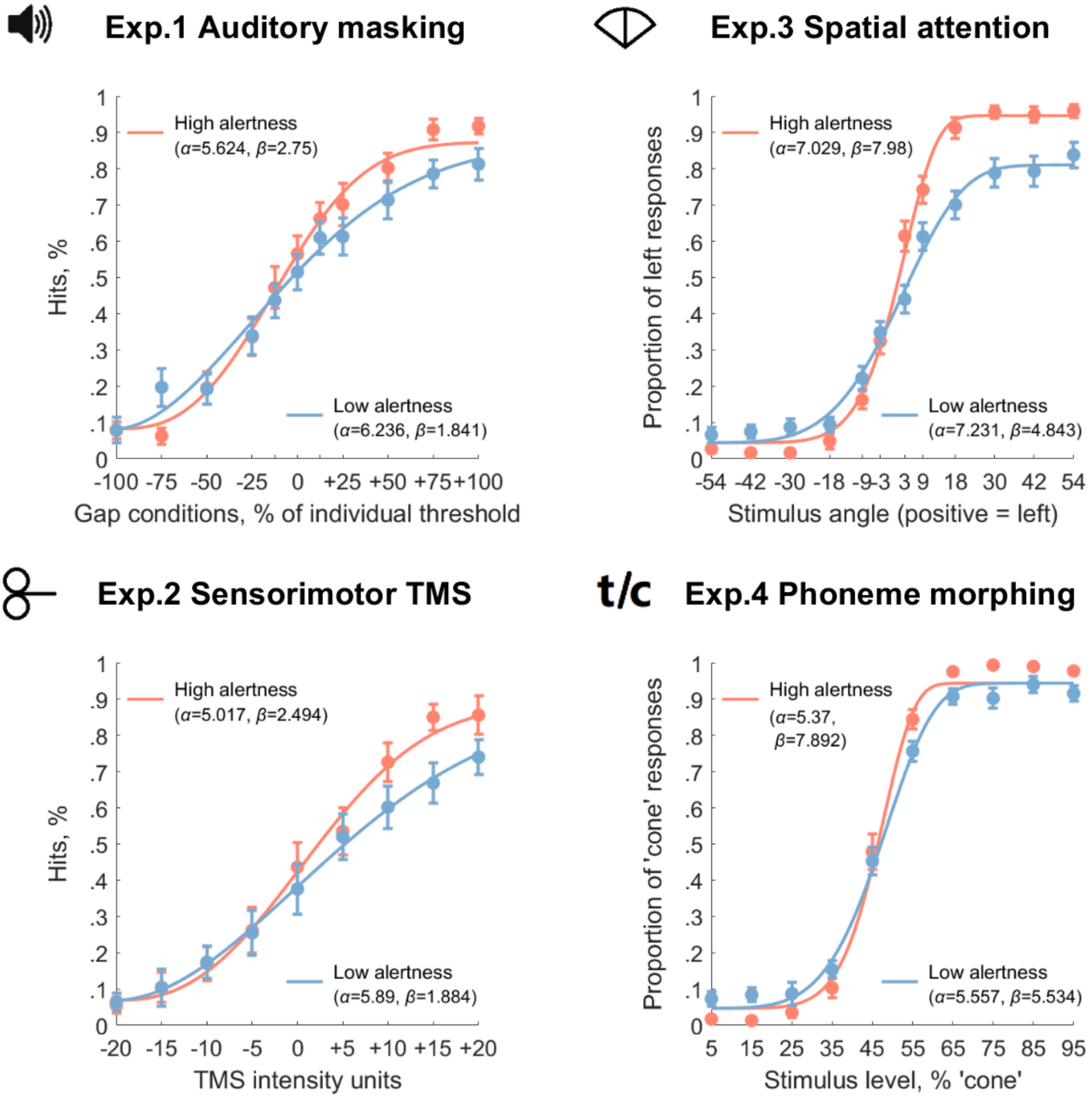

Fig. 2 – Psychometric curve fitting. High and low alertness sigmoidal curves are shown for all decision-making tasks, fitted with the Weibull function for each alertness level.[2] The error bars indicate variations across individual subjects. Alertness levels are defined by the micro-measures method. Slope (β) shows a systematic decrease with low alertness and the effect interacts with experiment settings (see Tables 1 and 2 for MLM results). Threshold (α) shows a reliable increase with lower alertness (Tables 3 and 4). See text for more details.

the change in stimulus intensity as they become drowsy. In other words, their precision of decision is reduced in lower alertness. To formally test this (hypothesis H2 in the preregistration), we compared the distributions of slopes between different alertness states for each experiment, and then performed MLM on the aggregated data across all tasks. The main finding is that alertness level has a strong effect on slope [$F(1, 80) = 51.228$, $p < .001$] as revealed by the analysis of variance in MLM (Table 2).

In detail, in the comparison of slope in each task, we found a medium to large effect of reduced slope for lower alertness level (Fig. 3). The experiment on spatial attention shows a larger effect than the other three experiments. Furthermore, slope shows a larger variance in the high alertness level in two tasks [$F(16, 16) = 3.193$, $p = .026$] for auditory masking task, $F(15, 15) = 2.546$, $p = .08$ for TMS task, $F(23, 23) = 2.075$, $p = .087$ for spatial attention task, and $F(22, 22) = 3.783$, $p = .003$ for

---

[2] We fitted an asymmetric function for the spatial attention task due to a leftward response bias (see Methods for details).

Table 1 – Model comparison for slope.

Model	Parameter	AIC	BIC	Log-likelihood	$p > (\chi^2)$
Null	Fixed: Mean, Random: Participant ID	894.39	903.62	−444.20	–
Alertness	Fixed: Alertness, Random: Participant ID	856.05	868.36	−424.03	2.14e-10[a]
Experiment	Fixed: Experiment, Random: Participant ID	812.75	831.20	−400.37	< 2.2e-16[a]
Alertness * Experiment	Fixed: Alertness * experiment, Random: Participant ID	757.23	787.98	−368.61	< 2.2e-16[a]

[a] $p < .001$.

Table 2 – Type III analysis of variance table for the alertness * experiment model on slope.

Model elements	Sum Sq	Mean Sq	NumDF	DenDF	F	p(>F)
Alertness	281.77	281.77	1	80	51.228	3.53e-10[a]
Experiment	990.15	330.05	3	80	60.006	< 2.2e-16[a]
Alertness:experiment	100.99	33.66	3	80	6.12	8.42e-04[a]

[a] $p < .001$.

Table 3 – Model comparison for threshold.

Model	Parameter	AIC	BIC	Log-likelihood	$p > (\chi^2)$
Null	Fixed: Mean, Random: Participant ID	481.19	490.45	−237.59	–
Alertness	Fixed: Alertness, Random: Participant ID	473.76	486.11	−232.88	.002[a]
Experiment	Fixed: Experiment, Random: Participant ID	425.69	444.21	−206.84	2.81e-13[b]
Alertness * Experiment	Fixed: Alertness * experiment, Random: Participant ID	420.07	450.94	−200.03	1.35e-13[b]

[a] $p < .01$.
[b] $p < .001$.

Table 4 – Type III analysis of variance table for the alertness * experiment model on threshold.

Model elements	Sum Sq	Mean Sq	NumDF	DenDF	F	p(>F)
Alertness	4.628	4.628	1	81	12.859	5.73e-04[a]
Experiment	33.139	11.046	3	81	30.692	2.36e-13[a]
Alertness:experiment	1.549	.516	3	81	1.435	.239

[a] $p < .001$.

phoneme discrimination task), pointing to a tendency to have larger individual differences in sensitivity when people are awake as compared to low alertness. To assess the impact of unequal variances, we employed robust standard errors and introduced a varIdent variance structure for the following MLM analysis on slope (see the supplementary section 3 for results supporting the robustness of our conclusions).

We further combined all tasks and performed MLM analysis using participant as a random factor and various combinations of alertness level and experiment as fixed factors to examine the effect of alertness level on slope among different task settings. The winning model (with the smallest values of AIC and BIC and the largest value of negative log-likelihood) included both alertness level and experiment factors (Table 1). As predicted in hypothesis H4, the main effect of alertness level on slope is significant (Table 2). This means that regardless of the different experimental settings, the sensitivity to stimulus intensity in perceptual decision-making is modulated by the alertness level of participants during the task. The factor ‘experiment’ alone also shows a reliable effect on slope [$F(3, 80) = 60.006$, $p < .001$]. As depicted in Figs. 2 and 3, the curves of discrimination tasks are steeper (larger slope) than that of the detection tasks. The spatial attention experiment exhibited an asymmetry for left and right stimuli, which is consistent with previous findings (Jagannathan et al., 2022).

We also observed a reliable interaction effect between alertness level and experiment [$F(3, 80) = 6.12$, $p < .001$]. The estimated decrease in slope of drowsiness varied across tasks,

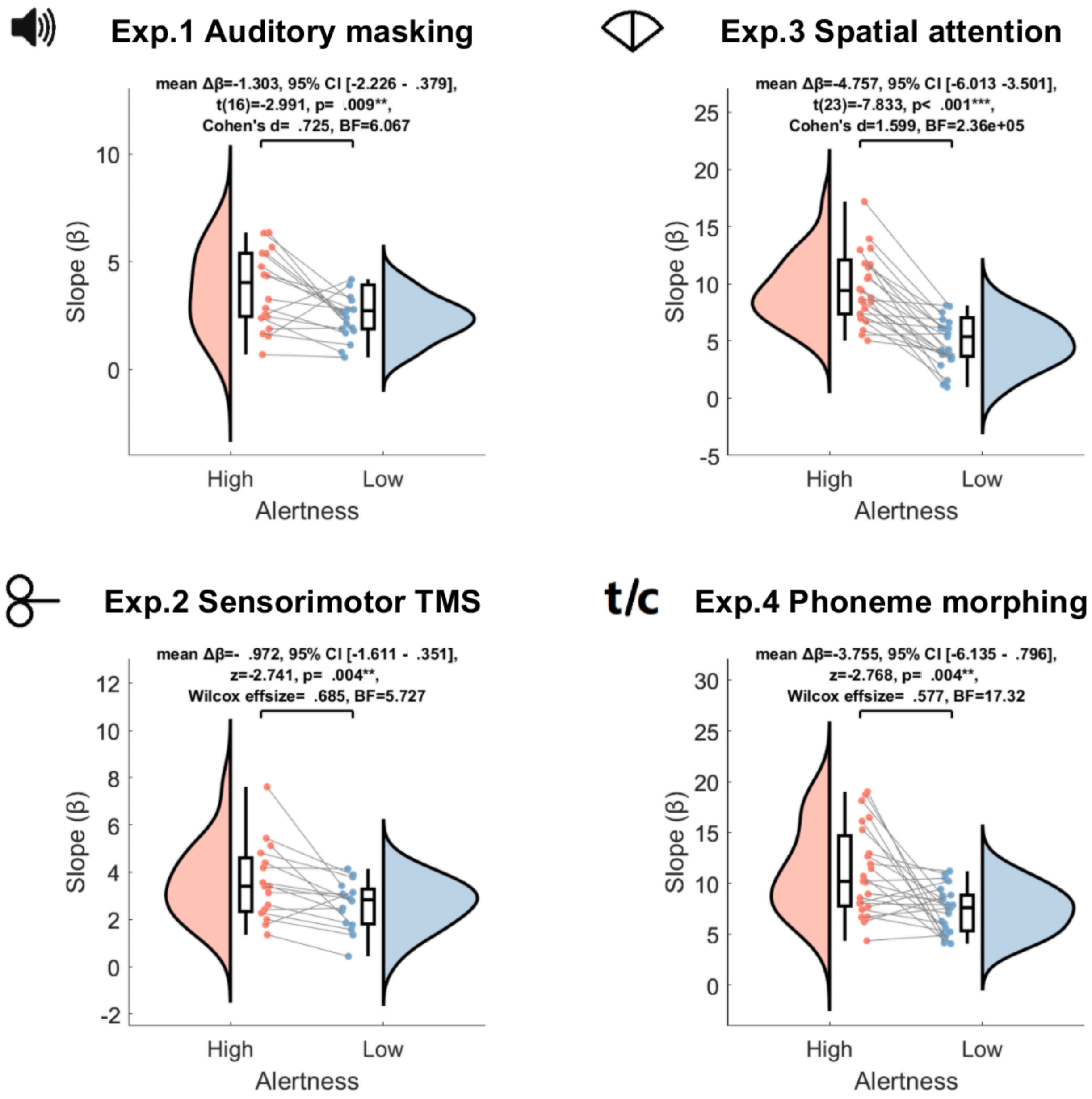

Fig. 3 – Distributions of slope. The distributions of slope are shown for each task and each alertness level (defined by the micro-measures method). Each data point represents the result from an individual participant. The results of the *t*-test or Wilcoxon signed-rank test, 95% confidence interval, effect size and the Bayes factor are shown in the figure. Slope decreases significantly with lower alertness level for all tasks. MLM reveals a reliable effect of alertness and an interaction effect between alertness and experiment settings (Tables 1 and 2). Post-hoc analysis shows that the decrease in slope with lower alertness is more pronounced in discrimination tasks ($p < .001$ for spatial attention and phoneme discrimination) than in detection tasks ($p = .737$ for auditory masking; $p = .94$ for the TMS task).

with larger effects observed in discrimination tasks (−4.757 for spatial attention, −3.755 for phoneme morphing) and smaller effects in detection tasks (−1.303 for auditory masking, −.972 for TMS-induced sensorimotor). Post-hoc analysis (Supplementary Table S3) confirmed that the effect of alertness level on slope was significant for discrimination tasks ($p < .001$ for both), but not for detection tasks ($p = .737$ and $p = .94$, respectively). These results not only reveal a general effect of alertness on the sensitivity of perceptual decision-making, but also show that the effect is stronger for discrimination than for detection.

To further test the results obtained from the micro-measures alertness split and look for convergent evidence, we performed the analysis using three other measures to categorise high and low alertness levels. These measures included the θ/α ratio, fast and slow RTs, and the proportion of missed responses within 10 trials. Alertness * experiment is the winning model in MLM analysis for all three ancillary methods as it was for the micro-measures method (Supplementary Table S1). The results confirm the findings with the micro-measures split method that the slope of the psychometric curve decreases as the alertness level gets lower, and the effect is reliable regardless of experiment type (Supplementary Table S2). We also find reliable effects of experiment and interaction on slope when using the other three split methods. Post-hoc analysis reveals that the larger effects of decreasing alertness for discrimination tasks are consistent among different split methods (Supplementary Table S3). Taken together, the main effects of low alertness and its interaction effects with experiment type on slope are convergent and robust across the methodologies used to define alertness.

Two complementary analyses were performed to confirm these results. First, we used the integrated session (high alert trials during the drowsy session) in the discrimination tasks to compare with the analysis using the awake session. The extra analysis showed similar results for the winning model and the effect of alertness, despite more lax conditions of high alertness levels (Supplementary section 4 Slope with drowsy sessions only). The interaction effects are, unsurprisingly, less consistent between methods, as interaction terms inherently require more data to estimate reliably than the main effects, making them more sensitive to differences in alertness classification. Second, to test the independence of our main effects from specific parameter specifications (guess rate γ and lapse rate λ) used to construct individual curves, we fitted the curve using a pair of group best γ and λ rates and a more strict pair ($\gamma = \lambda = .05$), respectively, showing that the individual variances in the curve fitting do not impair our main conclusion that alertness level has a common effect on slope across different experiments (Supplementary section 5 Slope with other fittings).

3.2. *Detection and discrimination thresholds increase with decreasing alertness*

We further characterise the effect of alertness on decision-making by testing its effects on threshold of the psychometric curve (hypothesis H3), revealing the limit of detection and discrimination ability and allowing us to see whether there is a bias in the participants' decision-making abilities with the change of alertness level. The comparison between the distributions of threshold in different states shows a small increase in threshold for low alertness (Fig. 4). Paired-sample t-tests show that the increase in threshold is substantial only for the TMS-induced kinaesthetic detection experiment ($t = 2.51$, $p = .027$, estimated threshold increase = .73).

To test the general effect of alertness on threshold, we also performed the MLM analysis as we did for slope. The alertness * experiment model has the smallest values of AIC and the largest value of negative log-likelihood, and the experiment model has the smallest values of BIC (Table 3). Further comparison shows that the alertness * experiment model is the winning model again ($\chi^2 = 13.621$, $p = .009$). The analysis of variance (Table 4) shows that the main effect of alertness is reliable [$F(1, 81) = 12.859$, $p < .001$], suggesting that threshold of the curves increases as the alertness level goes down as a common pattern of alertness as predicted in hypothesis H4. The factor ‘experiment’ also shows a main effect on threshold [$F(3, 81) = 30.692$, $p < .001$] while the interaction effect between alertness and experiment is not reliable [$F(3, 81) = 1.435$, $p = .239$]. To check the possibility that the overall effect of alertness is driven by TMS task alone, we also performed MLM on the other three tasks. Although the model with experiment as the only fixed factor is the best model, the analysis of variance for the alertness * experiment model still shows a reliable effect of alertness on threshold [$F(1, 68) = 5.032$, $p = .028$]. Therefore, we observe a common effect of alertness level on threshold of detection and discrimination decisions though it is not necessarily visible in each individual experiment.

Similar to slope, we also analysed threshold using the three other methods to split alertness. The results are mostly consistent with the micro-measures method: alertness * experiment is the winning model (Supplementary Table S12) for three methods except for omissions, in which there is no reliable evidence supporting an improvement of the alertness * experiment model compared to the experiment model ($\chi^2 = 8.976$, $p = .062$). Further ANOVA shows that both main effects of alertness and experiment are reliable across different methods (Supplementary Table S13). The only difference is that there is a reliable interaction effect under the RTs split [$F(3, 89) = 8.866$, $p < .001$], which comes from strong effects for the auditory masking task and TMS task as suggested by the post-hoc analysis (Supplementary Table S14). In addition, the complementary analysis using integrated sessions for alertness in the discrimination tasks yielded results consistent with those from the original analysis using separate sessions, showing that selecting awake trials from the alert session or from the drowsy session does not affect our findings for threshold (Supplementary section 7 Threshold with drowsy sessions only). Furthermore, we examined the influence of parameter specifications for the curve fitting, and found that the main effects of alertness and experiment are reliable for both group-best parameters and strict parameters, consistent with our main findings (Supplementary section 8 Threshold with other fittings). The interaction effects between alertness and experiment on threshold were found significant in three out of four splitting methods under both group-best and strict parameter settings (Supplementary Table S19). These effects are unlikely to reflect differences between detection and discrimination, but rather come from variations across individual experiments (Supplementary Table S20). The potential differential alertness effects between task types were suggested by the group fitting with the micro-measures method and the strict fitting with the RTs method; however, this pattern was not consistently observed across different parameter settings.

3.3. *Alertness affects perceptual sensitivity, but not criterion*

To further characterise the modulation exerted by fluctuations of alertness on perceptual decision-making, we applied SDT to complete the testing of hypotheses H2 and H3 defined in the preregistration. We assessed the effects of alertness on

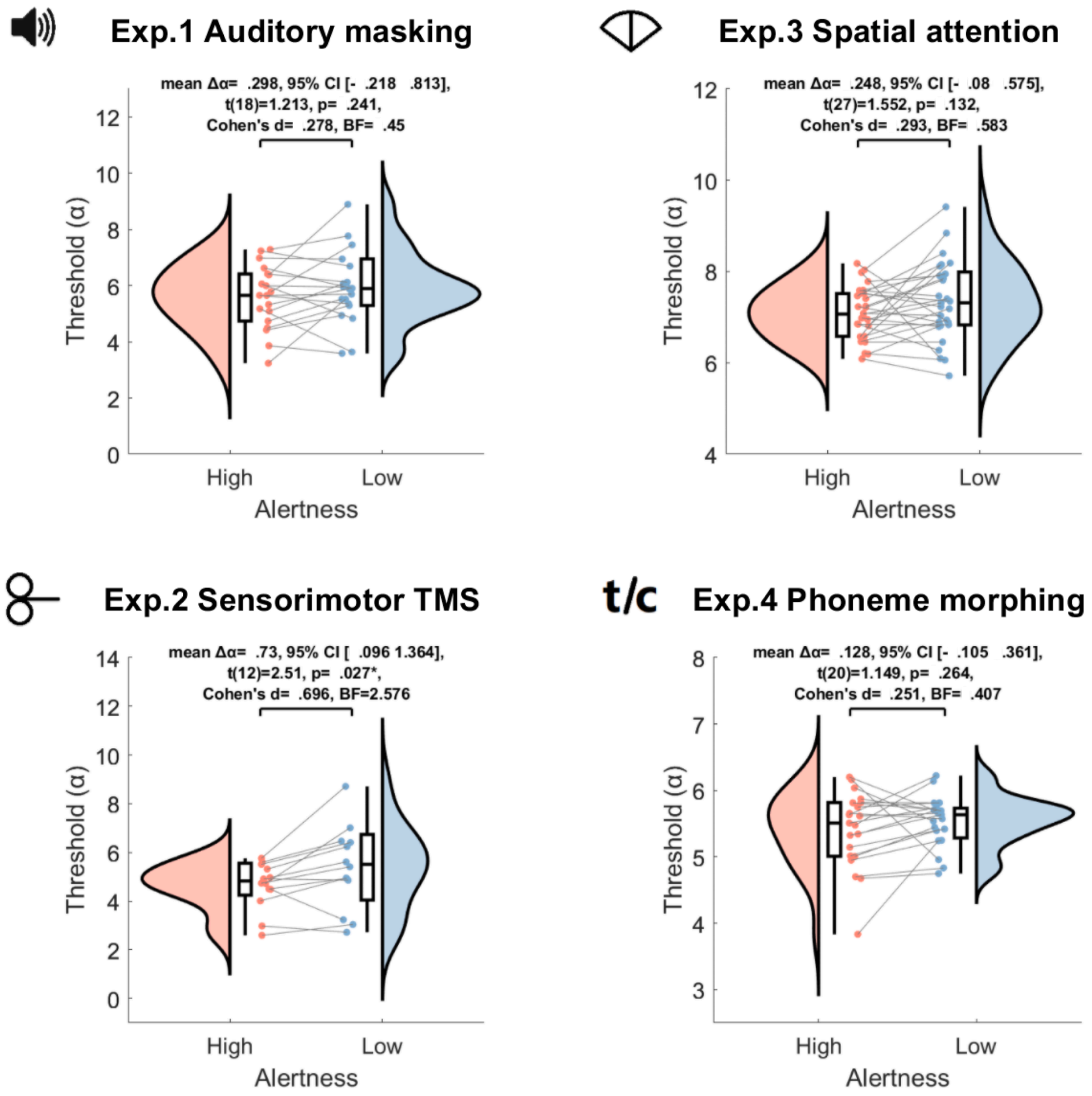

Fig. 4 – Distributions of threshold. Distributions of threshold are shown for each task and each alertness level (defined by the micro-measures method). Results of the *t*-test, 95% confidence interval, effect size and the Bayes factor are shown in the figure. Threshold shows a reliable increase in the TMS task and no obvious change in other tasks. MLM including all tasks reveals a reliable effect of alertness level on the threshold (see text and Tables 3 and 4).

perceptual sensitivity (d’) and the criterion (C) extracted from SDT analysis for each participant, and found that d’ decreases while C remains unchanged as the alertness level goes lower.

In detail, the comparison of the distribution of d’ is shown in Fig. 5. Compared to the high alertness level, the distributions of d’ tend to be lower and more variable in low alertness. There is a small effect in the detection tasks and a medium to large effect in the discrimination tasks. The variance of d’ increases in three tasks [$F(19, 19) = .293$, $p = .01$] for the auditory masking task, $F(26, 26) = .234$, $p < .001$ for the spatial attention task, and $F(32, 32) = .3$, $p = .001$ for the phoneme discrimination task) except for the TMS task [$F(15, 15) = .76$, $p = .602$]. In MLM, the alertness * experiment is the winning model (Table 5), and its ANOVA results show that both alertness [$F(1, 96) = 25.171$, $p < .001$] and experiment [$F(3, 96) = 8.846$, $p < .001$] have a strong effect on d’ (Table 6), indicating that there is a decrease in sensitivity as people get drowsy regardless of experiment types. There was no evidence for an interaction effect with d’ [$F(3, 96) = .88$, $p = .455$], which is different from what we found with slope of the sigmoidal curve. This implies that alertness

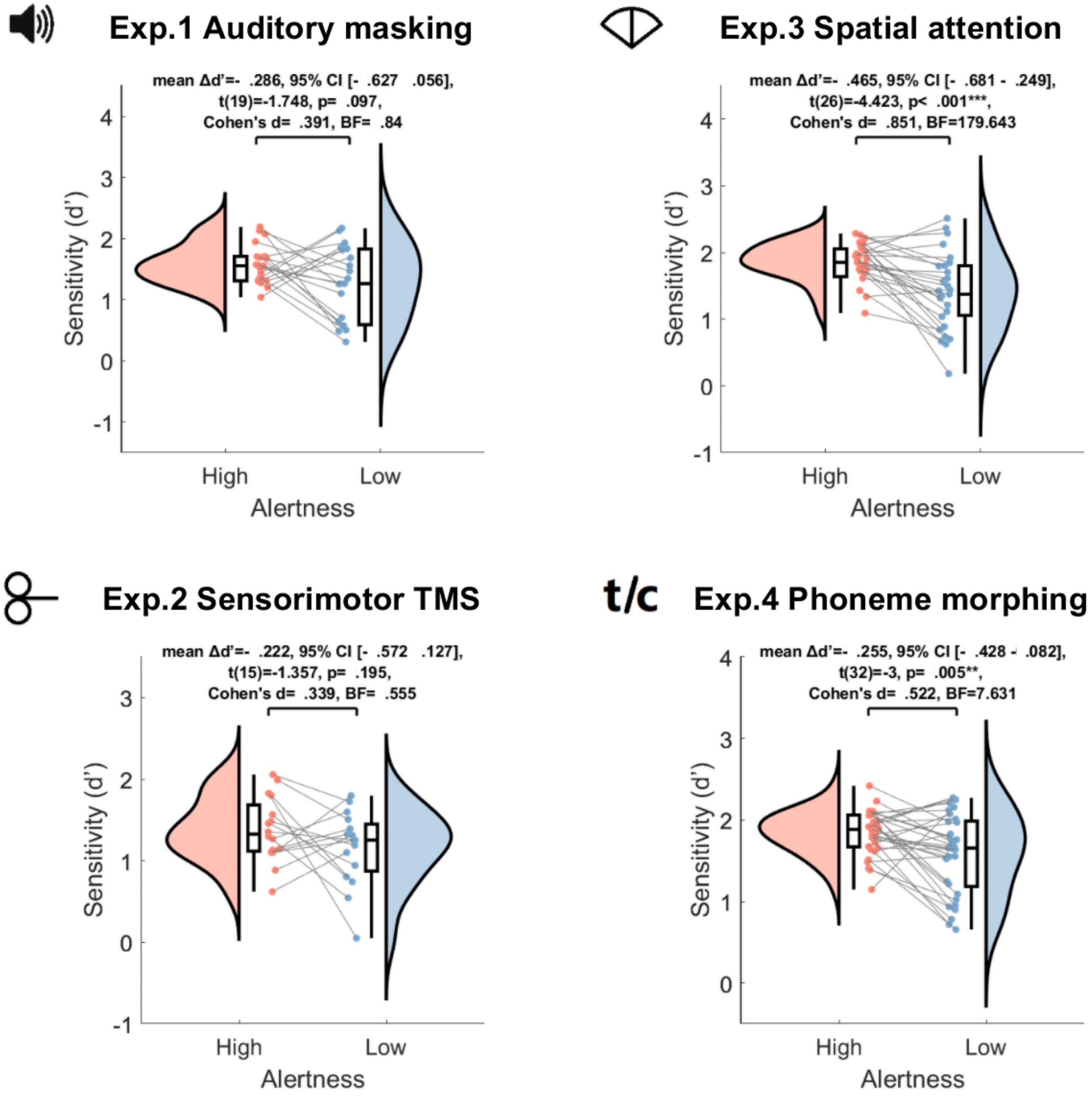

Fig. 5 – Distributions of d prime. The distributions of d' are shown for each task and each alertness level (defined by the micro-measures method). Results of the *t*-test, 95% confidence interval, effect size and the Bayes factor are shown in the figure. d' decreases reliably with alertness level for the spatial attention and phoneme discrimination tasks. Although there are no obvious effects in the detection tasks, the MLM including four experiments found a reliable effect of alertness and no interaction effect between alertness and experimental settings (see text and Tables 5 and 6).

Table 5 – Model comparison for d'.

Model	Parameter	AIC	BIC	Log-likelihood	$p > (\chi^2)$
Null	Fixed: Mean, Random: Participant ID	273.17	282.94	−133.59	–
Alertness	Fixed: Alertness, Random: Participant ID	250.73	263.76	−121.37	7.68e-07[a]
Experiment	Fixed: Experiment, Random: Participant ID	256.05	275.59	−122.02	3.81e-05[a]
Alertness * Experiment	Fixed: Alertness * experiment, Random: Participant ID	236.7	269.27	−108.35	1.17e-08[a]

[a] $p < .001$.

Table 6 – Type III analysis of variance table for the alertness * experiment model on d'.

Model elements	Sum Sq	Mean Sq	NumDF	DenDF	F	$p(>F)$
Alertness	4.189	4.189	1	96	25.171	2.41e-06[a]
Experiment	4.416	1.472	3	96	8.846	3.1e-05[a]
Alertness:experiment	.439	.146	3	96	.88	.455

[a] $p < .001$.

modulates detection and discrimination differentially when looking at how precise we are in making a decision but not for an overall measure of performance like perceptual sensitivity (d').

As with slope and threshold, the analyses of d' using the complementary categorising methods confirmed the convergence of the results obtained with the micro-measures split. Alertness * experiment is the winning model for all different splitting methods (Supplementary Table S21). Further ANOVA of the winning model (Supplementary Table S22) shows consistent outcomes with the micro-measures splitting method that alertness has a strong effect on d' while it shows no interaction with experiment for the RTs and omissions splitting methods. For the θ/α ratio method, there is a weak interaction effect between alertness and experiment on d' [$F(3, 107) = 2.707$, $p = .049$], which appears to be driven by the specific characteristics of the TMS experiment rather than differences between detection and discrimination tasks (Supplementary Table S23). The experiment factor also shows a reliable main effect on d' for all three complementary splitting methods. Taken together, these analyses confirm that d' decreases with lower alertness for all four alertness level categorisation methods, and show no evidence for an interaction effect between alertness and experiment type. Furthermore, the complementary analysis of integrated sessions also showed main effects of alertness and experiment on d', consistent with the original analysis. As for the interaction effects, we observed evidence in the micro-measures [$F(3, 90) = 3.305$, $p = .024$], θ/α ratio [$F(3, 111) = 5.513$, $p = .001$], and RTs [$F(3, 110) = 5.118$, $p = .002$] methods (Supplementary Table S25). Post-hoc tests showed that the significance comes from variations in individual experiments but can hardly be attributed to differences between detection and discrimination task types (Supplementary Table S26).

The other parameter of SDT, criterion, shows little changes from wakefulness to low alertness (Fig. 6). Results from MLM analysis show that the experiment model has the smallest values of AIC and BIC and the second largest log-likelihood value (Table 7). A comparison between the experiment model and the alertness * experiment model indicates that including alertness in the model does not significantly improve the fit ($\chi^2 = 2.404$, $p = .662$). ANOVA outcome of the alertness * experiment model (Table 8) does not suggest enough evidence for the effect of alertness [$F(1, 94) = .558$, $p = .457$] and for the interaction effect between alertness and experiment [$F(3, 94) = .438$, $p = .727$], while the main effect of experiment is reliable [$F(3, 94) = 10.323$, $p < .001$]. These results indicate that the decision-making criterion is unlikely to be influenced by alertness level.

The same analysis on criterion using θ/α splitting has consistent results as those from the micro-measures method (Supplementary Tables S27–S29). However, alertness shows a strong effect on criterion for the RTs and omissions methods (Supplementary Table S28). In addition to that, there is a reliable interaction effect [$F(3, 103) = 2.699$, $p = .05$] between alertness and experiment when using the RTs as the splitting method. Post-hoc analysis (Supplementary Table S29) reveals that criterion decreases as people enter a low alertness state for the spatial attention task when low alertness is defined by RTs or omissions, and for the phoneme morphing discrimination task when low alertness is defined by omissions. Thus, the alertness levels categorised by pre-stimulus brain state have little effect on criterion, whereas lower alertness defined by post-event behavioural markers significantly influences it. Additionally, the analysis using integrated sessions revealed a main effect of alertness [$F(1, 106) = 4.144$, $p = .044$] and an alertness × experiment interaction [$F(3, 106) = 3.043$, $p = .032$] in the θ/α ratio method (Supplementary Table S31), while post-hoc tests did not reveal sufficient evidence for alertness effects in individual experiments (Supplementary Table S32).

To summarise the results, the decision-making performance follows a sigmoidal curve in both high and low alertness states, but with different parameter characteristics of slope and threshold. The main findings show lower detection and discrimination sensitivity to stimuli in low alertness with a shallower slope of the psychometric curve and lower perceptual sensitivity (d'). This decrease in sensitivity is more pronounced for discrimination than for detection decisions as captured by the degree of change in the slope of the curve. Threshold shows a small increase with lower alertness and there is no change in the criterion to make the decision during fluctuations of alertness defined by pre-stimulus brain state.

4. Discussion

Despite the pervasiveness of alertness fluctuations occurring in every individual throughout the day, the underlying mechanisms by which they affect perceptual decision-making and cognitive functions are surprisingly understudied. Our study demonstrates an overwhelming agreement on how alertness fluctuations modulate perceptual decision-making. Here we developed an analysis framework using SDT and psychophysics to examine relative differences between high alertness and the metastable state of low alertness in two detection and two discrimination decision-making tasks. We showed that the slope of the sigmoidal curve decreases with alertness, associated with lower accuracy of sensory representation (Gold & Ding, 2013). Complementarily, we observed a subtle increase in threshold, indicating that participants may require more sensory evidence with lower alertness, although the increase is small and only evident in the

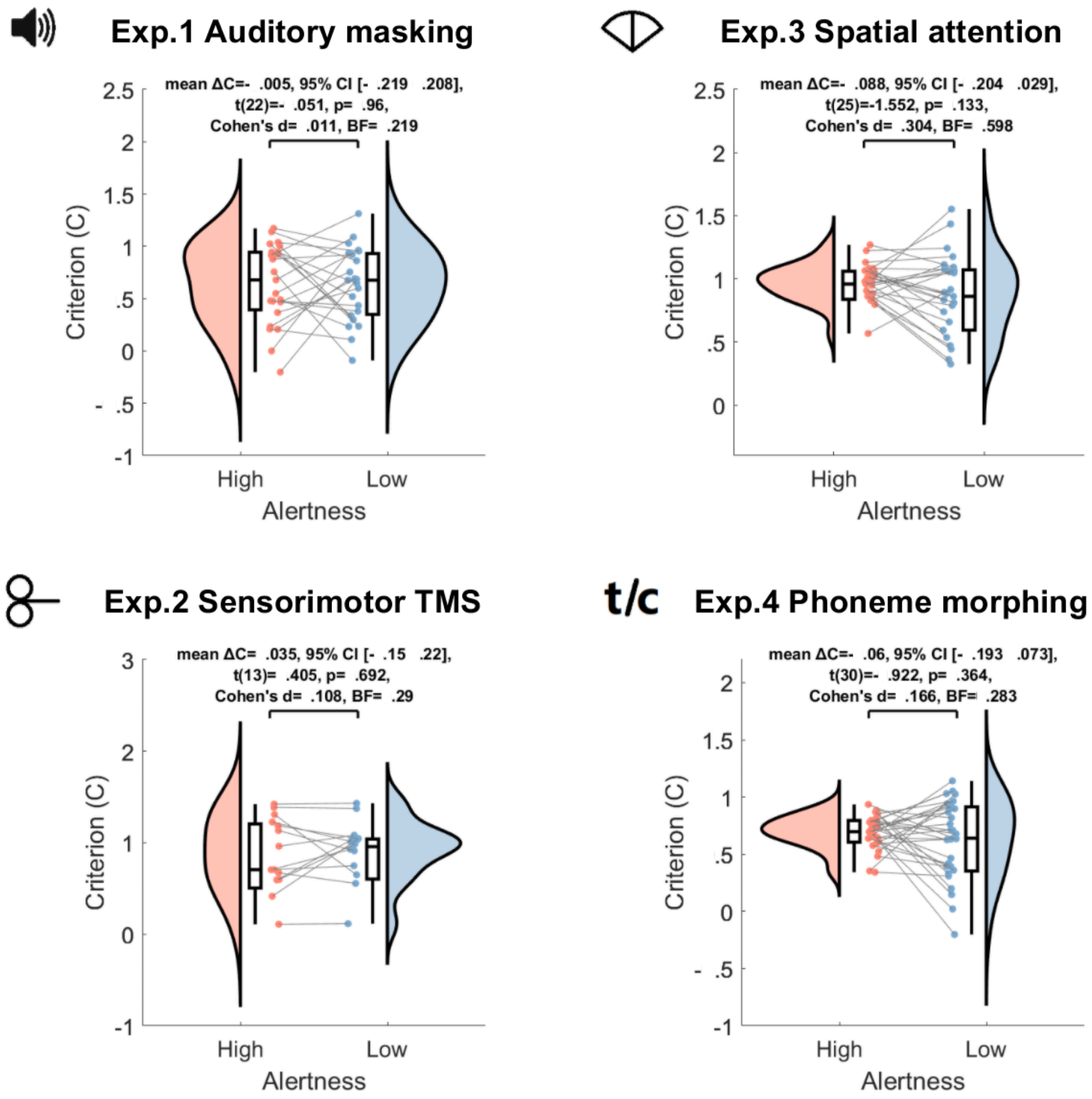

Fig. 6 – Distributions of criterion. Distributions of criterion are shown for each task and each alertness level (defined by the micro-measures method). Results of the *t*-test, 95% confidence interval, effect size and the Bayes factor are shown in the figure. Criterion shows no reliable change for all tasks.

Table 7 – Model comparison for C.

Model	Parameter	AIC	BIC	Log-likelihood	$p > (\chi^2)$
Null	Fixed: Mean, Random: Participant ID	105.5	115.21	−49.749	–
Alertness	Fixed: Alertness, Random: Participant ID	106.4	119.34	−49.199	.294
Experiment	Fixed: Experiment, Random: Participant ID	84.73	104.15	−36.365	6.58e-06[a]
Alertness * Experiment	Fixed: Alertness * experiment, Random: Participant ID	90.325	122.69	−35.163	1.35e-04[a]

[a] $p < .001$.

Table 8 – Type III analysis of variance table for the alertness * experiment model on C.

Model elements	Sum Sq	Mean Sq	NumDF	DenDF	F	$p(>F)$
Alertness	.038	.038	1	94	.558	.457
Experiment	2.098	.699	3	94	10.323	6.16e-06[a]
Alertness:experiment	.089	.03	3	94	.438	.727

[a] $p < .001$.

kinaesthetic detection experiment (TMS). In terms of general performance, d', capturing perceptual sensitivity, reveals a systematic and convergent decrease for all tasks, while the criterion, a reflection of internal rules to make a decision, may not vary. The changes observed in these four parameters, capturing complementary aspects of cognition, characterise the effects of alertness fluctuations on perceptual detection and discrimination in humans.

4.1. *Decreased decisions' precision with lower alertness may reflect a re-balance of the internal noise of the neural systems*

First, we found that performance in low alertness follows a similar sigmoidal psychometric curve as in the high alertness state but the curve is shallower (smaller slope) with decreasing alertness. Slope of the psychometric curve describes changes in response accuracy in relation to varying stimulus levels (Gold & Ding, 2013) and is often considered a reflection of the level of sensory noise (Morgan et al., 2012; Whiteley & Sahani, 2008). A decrease in slope with lower alertness was also observed in previous research using the same datasets of auditory detection, TMS-induced kinaesthetic detection and spatial discrimination that were analysed separately (Jagannathan et al., 2018; Noreika et al., 2017, 2020a). Here we found that the phenomenon is common across different experiment settings using a general analytical framework, indicating a diminished ability to precisely distinguish similar stimuli for both detection and discrimination tasks as people become less alert. This decreased precision of decision may reflect the increased variance in the underlying neural processes during the transition from wakefulness to lower alertness (Marreiros et al., 2008; Noreika et al., 2020a). In other words, people become less sensitive to exogenous noise (weak stimulus strength or introduction of another stimulus) as their endogenous noise increases (lower alertness level) in detection and discrimination decision-making. Importantly, external stimuli influence not only cognitive processes but also trigger phasic alertness fluctuations, which can in turn affect baseline alertness levels (tonic alertness). This influence occurs both transiently within the current trial and more enduringly for upcoming trials, by shaping internal states such as temporal expectations (Vangkilde et al., 2012) and urgency or time-pressure (Poth, 2021). Our findings suggest that reduced tonic alertness is more resilient to the counteracting effects of phasic alertness, revealing an interaction between alertness at multiple timescales.

The change in the precision of decision is also found in association with attention, which has been conceptualised as a measure of the inner noise of the system when integration of information is needed (Chennu et al., 2016; Lu & Dosher, 1998). Attention reduces the neural noise that is taken into account for the decision-making process, increasing the signal-to-noise ratio, and sensitivity as a result (Nunez et al., 2017; Smith & Ratcliff, 2009). Cohen and Maunsell (2011) studied how uncontrolled attentional fluctuations affect behavioural performance and found that the ability to detect minor changes in stimuli is considerably enhanced by attention. Herrmann et al. (2010) also found that attention could increase the slope of the psychometric function. Alertness and attention have an intertwined relationship. One key component of attention, sustained attention or vigilance, is not only integrated with cognitive processing but also connected to the alertness systems (Oken et al., 2006; Posner, 2008). The fluctuations of the neural dynamics underlying the perceptual and cognitive systems created by the transition of consciousness from high alertness to low alertness might show convergent effects to those proposed for sustained attention (Bareham et al., 2014), contributing in a similar manner to the changes in decision-making processes.

In addition, by combining four experiments together, here we surprisingly found a distinction showing that the decrease in slope of the psychometric curve accompanied by lower alertness was larger for discrimination than for detection decisions. Slope reflects the precision or sensitivity of detection around threshold in detection tasks and sensitivity to differences between stimuli in discrimination tasks (Prins, 2016). Thus the distinction indicates that fluctuations in the inner noise disproportionately impair sensitivity in discrimination decisions compared to detection. The interaction between noise and decision complexity has been proposed previously (Smith & Ratcliff, 2009). When external noise is introduced to the stimulus, its effect on performance is small in detection, but large in discrimination tasks (Smith et al., 2004; Smith & Wolfgang, 2007). Convergently, here we found an interaction between the inner noise (alertness level) and task types showing that the precision of decision declines more with low alertness for discrimination than detection. As far as the nature of the tasks is concerned, discrimination (distinguishing between two stimuli) has more decision complexity than detection (detecting one stimulus) (Sagi & Julesz, 1984; Smith & Ratcliff, 2009), and more complex neural networks may be involved in the former than in the latter (Correa et al., 2004). This makes discrimination easier to carry out as reflected by the steeper slope of the psychometric curve in discrimination than detection tasks. When alertness level declines, the increased inner noise may impair more information processing involved in discrimination as more neural networks may be affected, hence the decrease in the precision of decision (slope) is greater for discrimination than detection.

4.2. *Increased threshold with lower alertness may suggest compensation for less efficient information processing*

Next, the position of the psychometric curve is captured by threshold, which we found increased as alertness declines not systematically for each task but reliably in the aggregated data. When looking into the experiments individually, we found only a reliable change in threshold for the TMS sensorimotor task and not enough evidence of a shifted threshold for the other three tasks. This result confirms previous findings from the same datasets of auditory mask and TMS sensorimotor task (Noreika et al., 2017; 2020a). However, when combining data across all experiments, we could see a clear effect of alertness level on threshold regardless of the experiment type. The effect is not necessarily visible in every single experiment but needs aggregated data. Threshold represents the stimulus intensity needed for a certain performance level in detection tasks, or the strength of evidence at which one alternative is favored over the other in discrimination tasks (García-Pérez & Alcalá-Quintana, 2007; Macmillan & Creelman, 2005). Therefore, the observed increase in threshold suggests that participants either require stronger evidence to detect the target or show a shift in perceptual bias when discriminating between alternatives under lower alertness.

Threshold is also a reflection of the decision rule towards one alternative over another (Gold & Ding, 2013). Our result of the increased threshold indicates a propensity towards stronger target stimuli induced by a lower alertness level in decision-making tasks. Such a shift in decision rule has been considered as resulting from cognitive resource constraints such as time and memory capacity (Bossaerts & Murawski, 2017). Here the decreased alertness level results in cognitive and perceptual resource constraints and further prevents people from achieving optimal performance.

4.3. *Perceptual representations have broader and closer distributions with lower alertness*

Further analysis of SDT allows us to disentangle perceptual sensitivity and decision-making criteria during the transition from wakefulness to lower alertness. Our results show that perceptual sensitivity (d’) decreases as alertness level declines, which is consistent with previous research (Jagannathan et al., 2022). d’ reflects the underlying decision variable, typically modelled as a Gaussian distribution for each alternative in decision-making (Gold & Ding, 2013). A smaller d’ could indicate an increase in their common standard deviation and a decrease in the difference between the means of the Gaussian distributions. The former, increased standard deviation, indicates that inner noise becomes larger, making it more difficult to distinguish between the two. The latter represents that the perceptual representations of two alternatives are closer to each other which could also impair perceptual sensitivity. Either way, the decreased perception at sleep onset could be attributed to the diminished transmission of external information to the cortex, or its integration (Goupil & Bekinschtein, 2012).

Additionally, unlike the changes in slope, the variation of d’ is greater during low alertness across three tasks. It suggests that the perceptual sensitivity may become more dispersed among the population as people enter a lower alertness state, while the precision of decision (sensitivity to distinguish similar stimuli at the behavioural level) tends to be more convergent, for the group, in lower alertness compared to wakefulness. Furthermore, different from sensitivity to the stimuli (slope), the decrease in perceptual sensitivity with lower alertness shows no discrepancy between detection and discrimination tasks. It indicates that alertness level may comparably affect the sensitivity of detection and discrimination at the general performance level across different task conditions (measured by a normalised metric d’), but differentially at the local level around the threshold of the curve for difficult trials that are most indistinguishable (measured by slope).

On the other hand, we found no evidence of differences in SDT criterion across alertness levels, regardless of decision type. This result extends previous findings in the spatial attention task to wider decision-making scenarios (Jagannathan et al., 2022). Criterion in decision-making is a rule that the observer uses to partition the underlying distributions of stimuli (Macmillan & Creelman, 2005). The lack of evidence for a change in criterion in varied alertness may indicate a stable subjective rule while the sensory processing ability diminishes with lower alertness.

Based on the relation between SDT and the psychometric curve, we can see that the criterion to make decisions and the distance between representation distributions of the noise and the stimulus (or distributions of two stimuli in discrimination) contribute jointly to the threshold of performance. Since there is no evidence for a change in criterion at different alertness levels, the small increase in threshold we found should not be attributed to the change in criterion as we proposed in hypothesis H3. Instead, it could be caused by a decrease in the distance between distributions of alternatives with lower alertness. Therefore, there are indeed changes in response propensity (captured by threshold) due to changes in distribution distances which criterion (C) does not reflect. The smaller distance in lower alertness is also consistent with our finding of a decreased perceptual sensitivity (d’, computed as the difference in distance of distributions divided by the common standard deviation), although a small decrease in the distance could only account for part of the decline in sensitivity; the change in sensitivity also comes from the increased noise (standard deviation of the distribution) that is reported here.

4.4. *The effects of alertness on decision making are convergent between different alertness measures*

We did several variations of the analysis to validate our results. For the alertness level categorisation, we applied both EEG and behavioural measures to categorise alertness levels, to capture the different aspects of low alertness. Micromeasures method and θ/α ratio are both pretrial measures defined before the sensory perturbation and based on EEG/neural signals, reflecting a physiological ready state for decision-making (Samaha et al., 2020). On the other hand, split methods of RTs and proportion of omissions are measures of behaviour as the outcome of the perceptual decision (Goupil & Bekinschtein, 2012; Ogilvie et al., 1989). There is a consistency between pretrial neural measures and post-event behavioural measures, and both seem to point towards clear metastability of the arousal state *per se* and convergent on its modulation to cognitive performance except for criterion in SDT, as shown by the MLM results on aggregated data.

Notably, RT-based alertness effects could potentially be confounded with speed-accuracy trade-offs. While we acknowledge that the RTs measure alone cannot fully dissociate alertness effects from strategic speed-accuracy adjustments, it does not substantially impact our key findings as results are consistent when using non-RTs alertness measures. Moreover, our RTs method shows the opposite pattern of speed-accuracy trade-offs: low alertness (longer RTs) is associated with worse performance, whereas speed-accuracy trade-offs would predict improved accuracy with slower responses. This suggests that any undetected trade-off would likely weaken rather than explain our observed alertness-performance relationship, suggesting our results are more robust. Finally, analyses for Hypothesis 7 (Supplementary section 13) show that there is not enough evidence for a main effect of RT differences on changes in behavioural performance between high and low alertness, suggesting that our main findings are unlikely to be driven by reduced RT in lower alertness.

Alertness levels are fairly stable in neighbouring trials suggesting a clear hysteresis of the transition of consciousness even when perturbed systematically by tones, noise, words or tactile stimuli. Furthermore, previously our neural methods to separate different alertness states can only be applied when the participants have their eyes closed (Jagannathan et al., 2022). In current results, it seems that the type of perceptual decision does not influence the alertness metrics that we calculated, nor neural or behavioural, which is good news to extend the alertness-modulating cognition framework beyond auditory and tactile experiments. The convergence of effects of alertness on decision-making parameters seen between neural and behavioural methods to define high and low alertness periods is encouraging, and may help extend the separation between these states of the transition to other cases that involve visual or cross-modal integration since the behavioural methods are comparable even if the neural methods cannot be applied.

4.5. *Implications of the results*

The influence of alertness fluctuations on cognitive functions is insufficiently investigated despite its pervasiveness and wide application. This may be due to the historical separation between cognitive neuroscience and health and industrial-organizational (I/O) psychology (Butler & Senior, 2007). While cognitive neuroscience deals with the characterization of the underlying mechanisms of thought and action, applied psychology relates to human performance like driving behaviour and shift work. Applied psychology has been concentrating on capturing errors that can lead to accidents with little regard for neurocognitive mechanisms (Schreier et al., 2018). It is key to bring those willing to explain and those willing to predict real-life events together to create a truly translational programme of research that can use neurocognitive models to predict performance in a wide set of contexts and events. Here, we offer an example of the connection between performance and underlying cognitive factors in a common context of varying alertness levels, shedding light on both the explanation and prediction of perceptual decision-making.

Our convergent framework includes four methods for categorising alertness levels, capturing both pre-trial neural characteristics and post-event behavioural indicators, along with variations of psychometric curve-fitting specifications. The framework provides robust and reliable results of changes in decision-making performance in the transition from wakefulness to low alertness. In addition, the combination of four detection and discrimination experiments allows for testing the effect of alertness level across different modalities and experimental settings and offers a comparison between detection and discrimination decision-making. Furthermore, the joint application of psychometric function fitting and SDT connects behavioural performance to the underlying decision process, revealing that it is the broader and closer distributions of perceptual representation that lead to impaired performance of decision-making while the relative criterion remains stable from high alertness to low alertness. In this paper, this integrated analysis framework proves effective and promising for future use to illustrate how alertness level affects perceptual decision-making.

Our findings based on the integrated framework highlight the critical importance of measuring and controlling alertness-related noise in experimental designs, particularly for tasks involving perceptual discrimination. Our results show that EEG-based pre-trial measures can effectively predict alertness fluctuations, allowing either (a) proactive exclusion of drowsy trials or (b) statistical accounting for alertness effects. This approach enhances the reliability of the data by minimising the influence of alertness on task performance. Furthermore, the observed dissociation between stable decision criteria and altered perceptual processing indicates that alertness effects primarily influence perceptual sensitivity rather than decision strategy. These effects seem task dependent, as strategy is severely impacted in more complex tasks like reversal learning and other forms of cognitive control (Alameda et al., 2024; Ciria et al., 2021). These findings underscore the differential impacts of alertness fluctuations across cognitive domains, which should be carefully considered in experimental design.

Moving forward, this research opens up new avenues for future investigation. To start, the relationship among different methods of alertness classification could be further explored. Here we found convergence regarding performance outcomes, and it would be informative to see the degree of agreement these methods show. Considerable overlap of alertness categorisation is expected, and the features of those most consistent trials could be further explored with the potential to develop an integrated and comprehensive classification method for alertness levels. Further, current evidence of increased internal noise and demand for stronger stimuli at lower alertness levels suggests changes in information processing during the transition between alertness levels. Future works would benefit from applying information theory tools to explore the underlying information dynamics of decision-making in fluctuating alertness.

CRediT authorship contribution statement

Yanzhi Xu: Writing – original draft, Formal analysis, Data curation. **Martijn Wokke:** Writing – review & editing, Supervision, Formal analysis, Conceptualization. **Valdas Noreika:** Writing – review & editing, Methodology, Data curation. **Corinne Bareham:** Writing – review & editing, Methodology, Data curation. **Sridhar Jagannathan:** Writing – review & editing, Methodology, Data curation. **Stanimira Georgieva:** Writing – review & editing, Methodology, Data curation. **Caterina Trentin:** Writing – review & editing, Methodology, Data curation. **Tristan Bekinschtein:** Writing – review & editing, Supervision, Conceptualization.

Scientific transparency statement

DATA: Some raw and processed data supporting this research are publicly available, while some are subject to restrictions: https://doi.org/10.17863/CAM.95292, https://doi.org/10.5281/zenodo.15783234, https://doi.org/10.5281/zenodo.5655444, https://doi.org/10.5281/zenodo.15763453.

CODE: All analysis code supporting this research is publicly available: https://doi.org/10.17863/CAM.95292.

MATERIALS: This research did not make use of any materials to generate or acquire data.

DESIGN: This article reports, for all studies, how the author(s) determined all sample sizes, all data exclusions, all data inclusion and exclusion criteria, and whether inclusion and exclusion criteria were established prior to data analysis.

PRE-REGISTRATION: No part of the study procedures was pre-registered in a time-stamped, institutional registry prior to the research being conducted. At least part of the analysis plans was pre-registered in a time-stamped, institutional registry prior to the research being conducted: https://osf.io/3p4rf The analyses that were undertaken deviated from the preregistered analysis plans. All such deviations are fully disclosed in the manuscript.

For full details, see the *Scientific Transparency Report* in the supplementary data to the online version of this article.

Funding

This work was supported by China Scholarship Council and Cambridge Trust (201906010293 to Yanzhi Xu) and the Wellcome Trust (WT093811MA to Tristan Bekinschtein). The authors have no conflicts of interest to declare.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2025.06.018.

REFERENCES

Alameda, C., Avancini, C., Sanabria, D., Bekinschtein, T. A., Canales-Johnson, A., & Ciria, L. F. (2024). Staying in control: Characterizing the mechanisms underlying cognitive control in high and low arousal states. *British Journal of Psychology, 115*(4), 665–682. https://doi.org/10.1111/bjop.12715

Allen, M., Poggiali, D., Whitaker, K., Marshall, T., van Langen, J., & Kievit, R. (2021). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research, 4*(63). https://doi.org/10.12688/wellcomeopenres.15191.2

Bareham, C. A., Bekinschtein, T. A., Scott, S. K., & Manly, T. (2015). Does left-handedness confer resistance to spatial bias? *Scientific Reports, 5*(1), 9162. https://doi.org/10.1038/srep09162

Bareham, C. A., Manly, T., Pustovaya, O. V., Scott, S. K., & Bekinschtein, T. A. (2014). Losing the left side of the world: Rightward shift in human spatial attention with sleep onset. *Scientific Reports, 4*(1), 1–5. https://doi.org/10.1038/srep05092

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, M. B. (2015). Package 'lme4'. https://www.r-project.org/package=lme4.

Baulk, S. D., Reyner, L., & Horne, J. A. (2001). Driver sleepiness—evaluation of reaction time measurement as a secondary task. *Sleep, 24*(6), 695–698. https://doi.org/10.1093/sleep/24.6.695

Bekinschtein, T., Cologan, V., Dahmen, B., & Golombek, D. (2009). You are only coming through in waves: Wakefulness variability and assessment in patients with impaired consciousness. *Progress in Brain Research, 177*, 171–189. https://doi.org/10.1016/S0079-6123(09)17712-9

Borbély, A. A., Daan, S., Wirz-Justice, A., & Deboer, T. (2016). The two-process model of sleep regulation: A reappraisal. *Journal of Sleep Research, 25*(2), 131–143. https://doi.org/10.1111/jsr.12371

Bossaerts, P., & Murawski, C. (2017). Computational complexity and human decision-making. *Trends in Cognitive Sciences, 21*(12), 917–929. https://doi.org/10.1016/j.tics.2017.09.005

Bridwell, D. A., Hecker, E. A., Serences, J. T., & Srinivasan, R. (2013). Individual differences in attention strategies during detection, fine discrimination, and coarse discrimination. *Journal of Neurophysiology, 110*(3), 784–794. https://doi.org/10.1152/jn.00520.2012

Butler, M. J., & Senior, C. (2007). Toward an organizational cognitive neuroscience. *Annals of the New York Academy of Sciences, 1118*(1), 1–17. https://doi.org/10.1196/annals.1412.009

Canales-Johnson, A., Beerendonk, L., Blain, S., Kitaoka, S., Ezquerro-Nassar, A., Nuiten, S., Fahrenfort, J., van Gaal, S., & Bekinschtein, T. A. (2020). Decreased alertness reconfigures cognitive control networks. *Journal of Neuroscience, 40*(37), 7142–7154. https://doi.org/10.1523/JNEUROSCI.0343-20.2020

Carrier, J., & Monk, T. H. (2000). Circadian rhythms of performance: New trends. *Chronobiology International, 17*(6), 719–732. https://doi.org/10.1081/CBI-100102108

Chennu, S., Noreika, V., Gueorguiev, D., Shtyrov, Y., Bekinschtein, T. A., & Henson, R. (2016). Silent expectations: Dynamic causal modeling of cortical prediction and attention to sounds that weren't. *Journal of Neuroscience, 36*(32), 8305–8316. https://doi.org/10.1523/JNEUROSCI.1125-16.2016

Ciria, L. F., Suárez-Pinilla, M., Williams, A. G., Jagannathan, S. R., Sanabria, D., & Bekinschtein, T. A. (2021). Different underlying mechanisms for high and low arousal in probabilistic learning in humans. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 143*, 180–194. https://doi.org/10.1016/j.cortex.2021.07.002

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. The MIT press. https://doi.org/10.7551/mitpress/9609.001.0001

Cohen, M. R., & Maunsell, J. H. (2011). When attention wanders: How uncontrolled fluctuations in attention affect performance. *Journal of Neuroscience, 31*(44), 15802–15806. https://doi.org/10.1523/JNEUROSCI.3063-11.2011

Comsa, I. M., Bekinschtein, T. A., & Chennu, S. (2019). Transient topographical dynamics of the electroencephalogram predict brain connectivity and behavioural responsiveness during drowsiness. *Brain Topography, 32*(2), 315–331. https://doi.org/10.1007/s10548-018-0689-9

Correa, Á., Lupiáñez, J., Milliken, B., & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Perception & Psychophysics, 66*(2), 264–278. https://doi.org/10.3758/BF03194878

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, 5(10), Article e260. https://doi.org/10.1371/journal.pbio.0050260

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

Diaz, B. A., Hardstone, R., Mansvelder, H. D., Van Someren, E. J., & Linkenkaer-Hansen, K. (2016). Resting-state subjective experience and eeg biomarkers are associated with sleep-onset latency. *Frontiers in Psychology, 7*, 492. https://doi.org/10.3389/fpsyg.2016.00492

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex, 23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences, 1396*(1), 70–91. https://doi.org/10.1111/nyas.13318

García-Pérez, M. A., & Alcalá-Quintana, R. (2007). The transducer model for contrast detection and discrimination: Formal relations, implications, and an empirical test. *Spatial Vision, 20*. https://doi.org/10.1163/156856807779369724

Goel, N., Van Dongen, H. P., & Dinges, D. F. (2011). Circadian rhythms in sleepiness, alertness, and performance. In *Principles and practice of sleep medicine* (pp. 445–455). Elsevier. https://doi.org/10.1016/B978-1-4160-6645-3.00038-4.

Gold, J. I., & Ding, L. (2013). How mechanisms of perceptual decision-making affect the psychometric function. *Progress in Neurobiology, 103*, 98–114. https://doi.org/10.1016/j.pneurobio.2012.05.008

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

Goupil, L., & Bekinschtein, T. (2012). Cognitive processing during the transition to sleep. *Archives Italiennes de Biologie, 150*(2), 140–154. https://doi.org/10.4449/aib.v150i2.1247

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). John Wiley.

Hancock, P. A. (2013). In search of vigilance: The problem of iatrogenically created psychological phenomena. *American Psychologist, 68*(2), 97–109. https://doi.org/10.1037/a0030214

Hanks, T. D., & Summerfield, C. (2017). Perceptual decision making in rodents, monkeys, and humans. *Neuron, 93*(1), 15–31. https://doi.org/10.1016/j.neuron.2016.12.003

Harrower, M., & Brewer, C. A. (2003). Colorbrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal, 40*(1), 27–37. https://doi.org/10.1179/000870403235002042

Harsh, J., Voss, U., Hull, J., Schrepfer, S., & Badia, P. (1994). ERP and behavioral changes during the wake/sleep transition. *Psychophysiology, 31*(3), 244–252. https://doi.org/10.1111/j.1469-8986.1994.tb02213.x

Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., & Heeger, D. J. (2010). When size matters: Attention affects performance by contrast or response gain. *Nature Neuroscience, 13*(12), 1554–1559. https://doi.org/10.1038/nn.2669

Hori, T., Hayashi, M., & Morikawa, T. (1994). Topographical EEG changes and the hypnagogic experience. In *Sleep onset: Normal and abnormal processes* (pp. 237–253). https://doi.org/10.1037/10166-014

Ikoma, K., Samii, A., Mercuri, B., Wassermann, E., & Hallett, M. (1996). Abnormal cortical motor excitability in dystonia. *Neurology, 46*(5), 1371. https://doi.org/10.1212/WNL.46.5.1371, 1371.

Jagannathan, S. R., Bareham, C. A., & Bekinschtein, T. A. (2022). Decreasing alertness modulates perceptual decision-making. *Journal of Neuroscience, 42*(3), 454–473. https://doi.org/10.1523/JNEUROSCI.0182-21.2021

Jagannathan, S. R., Ezquerro-Nassar, A., Jachs, B., Pustovaya, O. V., Bareham, C. A., & Bekinschtein, T. A. (2018). Tracking wakefulness as it fades: Micro-measures of alertness. *Neuroimage, 176*, 138–151. https://doi.org/10.1016/j.neuroimage.2018.04.046

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The epworth sleepiness scale. *Sleep, 14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

Jung, T. P., Makeig, S., Stensmo, M., & Sejnowski, T. J. (1997). Estimating alertness from the eeg power spectrum. *IEEE Transactions on Biomedical Engineering, 44*(1), 60–69. https://doi.org/10.1109/10.553713

Koch, C., & Preuschoff, K. (2007). Betting the house on consciousness. *Nature Neuroscience, 10*(2), 140–141. https://doi.org/10.1038/nn0207-140

Lagarde, D., & Batejat, D. (1994). Evaluation of drowsiness during prolonged sleep deprivation. *Neurophysiologie Clinique/Clinical Neurophysiology, 24*(1), 35–44. https://doi.org/10.1016/S0987-7053(05)80403-4

Liberson, W., & Liberson, C. W. (1966). EEG records, reaction times, eye movements, respiration, and mental content during drowsiness. In *Recent advances in biological psychiatry* (pp. 295–302). Springer. https://doi.org/10.1007/978-1-4899-7313-9_31.

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*(9), 1183–1198. https://doi.org/10.1016/S0042-6989(97)00273-3

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). New York. NY: Lawrence Erlbaum Associates Publishers.

Makeig, S., Jung, T. P., & Sejnowski, T. J. (2000). Awareness during drowsiness: Dynamics and electrophysiological correlates. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 54*(4), 266. https://doi.org/10.1037/h0087346

Marcel, A. J. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology, 15*(2), 197–237. https://doi.org/10.1016/0010-0285(83)90009-9

Marreiros, A. C., Daunizeau, J., Kiebel, S. J., & Friston, K. J. (2008). Population dynamics: Variance and the sigmoid activation function. *Neuroimage, 42*(1), 147–157. https://doi.org/10.1016/j.neuroimage.2008.04.239

McGinley, M. J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C. R., Tolias, A. S., Cardin, J. A., & McCormick, D. A. (2015). Waking state: Rapid variations modulate neural and behavioral responses. *Neuron, 87*(6), 1143–1161. https://doi.org/10.1016/j.neuron.2015.09.012

Mediano, P. A., Ikkala, A., Kievit, R. A., Jagannathan, S. R., Varley, T. F., Stamatakis, E. A., Bekinschtein, T. A., & Bor, D. (2021). Fluctuations in neural complexity during wakefulness relate to conscious level and cognition. *BioRxiv*. https://doi.org/10.1101/2021.09.23.461002

Moller, H. J., Devins, G. M., Shen, J., & Shapiro, C. M. (2006). Sleepiness is not the inverse of alertness: Evidence from four sleep disorder patient groups. *Experimental Brain Research, 173*, 258–266. https://doi.org/10.1007/s00221-006-0436-4

Morgan, M., Dillenburger, B., Raphael, S., & Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity. *Attention, Perception, & Psychophysics, 74*, 185–193. https://doi.org/10.3758/s13414-011-0222-7

Noreika, V., Canales-Johnson, A., Johnson, A., Arnatkevičiūtė, A., Koh, J., Chennu, S., & Bekinschtein, T. A. (2020a). Wakefulness fluctuations elicit behavioural and neural reconfiguration of awareness. *bioRxiv*. https://doi.org/10.1101/155705

Noreika, V., Kamke, M. R., Canales-Johnson, A., Chennu, S., Bekinschtein, T. A., & Mattingley, J. B. (2020b). Alertness fluctuations when performing a task modulate cortical evoked responses to transcranial magnetic stimulation. *Neuroimage, 223*, Article 117305. https://doi.org/10.1016/j.neuroimage.2020.117305

Noreika, V., Kamke, M. R., Canales-Johnson, A., Chennu, S., Mattingley, J. B., & Bekinschtein, T. A. (2017). Neurobehavioral dynamics of drowsiness. *bioRxiv*, 1–36. https://doi.org/10.1101/155754

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

O'Connell, R. G., Shadlen, M. N., Wong-Lin, K., & Kelly, S. P. (2018). Bridging neural and computational viewpoints on perceptual decision-making. *Trends in Neurosciences, 41*(11), 838–852. https://doi.org/10.1016/j.tins.2018.06.005

Ogilvie, R. D. (2001). The process of falling asleep. *Sleep Medicine Reviews, 5*(3), 247–270. https://doi.org/10.1053/smrv.2001.0145

Ogilvie, R. D., Simons, I. A., Kuderian, R. H., MacDonald, T., & Rustenburg, J. (1991). Behavioral, event-related potential, and EEG/FFT changes at sleep onset. *Psychophysiology, 28*(1), 54–64. https://doi.org/10.1111/j.1469-8986.1991.tb03386.x

Ogilvie, R. D., Wilkinson, R. T., & Allison, S. (1989). The detection of sleep onset: Behavioral, physiological, and subjective convergence. *Sleep, 12*(5), 458–474. https://doi.org/10.1093/sleep/12.5.458

Oken, B. S., Salinsky, M. C., & Elsas, S. (2006). Vigilance, alertness, or sustained attention: Physiological basis and measurement. *Clinical Neurophysiology, 117*(9), 1885–1901. https://doi.org/10.1016/j.clinph.2006.01.017

Posner, M. I. (2008). Measuring alertness. *Annals of the New York Academy of Sciences, 1129*(1), 193–199. https://doi.org/10.1196/annals.1417.011

Poth, C. H. (2021). Urgency forces stimulus-driven action by overcoming cognitive control. *Elife, 10*, Article e73682. https://doi.org/10.7554/eLife.73682

Poth, C. H. (2025). Readiness for perception and action: Towards a more mechanistic understanding of phasic alertness. *Journal of Cognition, 8*(1), 19. https://doi.org/10.5334/joc.426

Prins, N. (2016). *Psychophysics: A practical introduction*. Academic Press. https://doi.org/10.1016/C2012-0-01278-1

Prins, N., & Kingdom, F. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the palamedes toolbox. *Frontiers in Psychology, 9*, 1250. https://doi.org/10.3389/fpsyg.2018.01250

R Core Team. (2013). R: A language and environment for statistical computing. https://www.r-project.org/.

Rossini, P. M., Barker, A. T., Berardelli, A., Caramia, M. D., Caruso, G., Cracco, R. Q., Dimitrijević, M. R., Hallett, M., Katayama, Y., Lücking, C. H., Noordhout, A. L. M.d., Marsden, C. D., Murray, N. M. F., Rothwell, J. C., Swash, M., & Tomberg, C. (1994). Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: Basic principles and procedures for routine clinical application. report of an IFCN committee. *Electroencephalography and Clinical Neurophysiology, 91*(2), 79–92. https://doi.org/10.1016/0013-4694(94)90029-9

Sagi, D., & Julesz, B. (1984). Detection versus discrimination of visual orientation. *Perception, 13*(5), 619–628. https://doi.org/10.1068/p130619

Samaha, J., Iemi, L., Haegens, S., & Busch, N. A. (2020). Spontaneous brain oscillations and perceptual decision-making. *Trends in Cognitive Sciences, 24*(8), 639–653. https://doi.org/10.1016/j.tics.2020.05.004

Samii, A., Wassermann, E., Ikoma, K., Mercuri, B., & Hallett, M. (1996). Characterization of postexercise facilitation and depression of motor evoked potentials to transcranial magnetic stimulation. *Neurology*, 46(5), 1376. https://doi.org/10.1212/WNL.46.5.1376, 1376.

Sandberg, K., Bibby, B. M., Timmermans, B., Cleeremans, A., & Overgaard, M. (2011). Measuring consciousness: Task accuracy and awareness as sigmoid functions of stimulus duration. *Consciousness and Cognition, 20*(4), 1659–1675. https://doi.org/10.1016/j.concog.2011.09.002

Schreier, D. R., Banks, C., & Mathis, J. (2018). Driving simulators in the clinical assessment of fitness to drive in sleepy individuals: A systematic review. *Sleep Medicine Reviews, 38*, 86–100. https://doi.org/10.1016/j.smrv.2017.04.004

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science, 15*(11), 720–728. https://doi.org/10.1111/j.0956-7976.2004.00748.x

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

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116*(2), 283. https://doi.org/10.1037/a0015156

Smith, P. L., & Wolfgang, B. J. (2007). Attentional mechanisms in visual signal detection: The effects of simultaneous and delayed noise and pattern masks. *Perception & Psychophysics, 69*(7), 1093–1104. https://doi.org/10.3758/BF03193947

Smith, P. L., Wolfgang, B. J., & Sinclair, A. J. (2004). Mask-dependent attentional cuing effects in visual signal detection: The psychometric function for contrast. *Perception & Psychophysics, 66*(6), 1056–1075. https://doi.org/10.3758/BF03194995

Steghaus, S., & Poth, C. H. (2024). Feeling tired versus feeling relaxed: Two faces of low physiological arousal. *Plos One, 19*(9), Article e0310034. https://doi.org/10.1371/journal.pone.0310034

Thomas, J. P., & Gille, J. (1979). Bandwidths of orientation channels in human vision. *JOSA, 69*(5), 652–660. https://doi.org/10.1364/JOSA.69.000652

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience, 16*, 601–615. https://doi.org/10.3758/s13415-016-0417-4

van Kempen, J., Loughnane, G. M., Newman, D. P., Kelly, S. P., Thiele, A., O'Connell, R. G., & Bellgrove, M. A. (2019). Behavioural and neural signatures of perceptual decision-making are modulated by pupil-linked arousal. *eLife, 8*, Article e42541. https://doi.org/10.7554/eLife.42541

Vangkilde, S., Coull, J. T., & Bundesen, C. (2012). Great expectations: Temporal expectation modulates perceptual processing speed. *Journal of Experimental Psychology: Human Perception and Performance, 38*(5), 1183. https://doi.org/10.1037/a0026343

Waschke, L., Tune, S., & Obleser, J. (2019). Local cortical desynchronization and pupil-linked arousal differentially shape brain states for optimal sensory performance. *eLife, 8*, Article e51501. https://doi.org/10.7554/eLife.51501

Whiteley, L., & Sahani, M. (2008). Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *Journal of Vision, 8*(3), 2. https://doi.org/10.1167/8.3.2, 2.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & psychophysics, 63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Williams, H. L., Lubin, A., & Goodnow, J. J. (1959). Impaired performance with acute sleep loss. *Psychological Monographs: General and Applied, 73*(14), 1. https://doi.org/10.1037/h0093749

Xu, Y., Wokke, M., Noreika, V., Bareham, C., Jagannathan, S. R., Georgieva, S., Trentin, C., & Bekinschtein, T. A. (2023). Effects of alertness on perceptual detection and discrimination. *BioRxiv*, 2023. https://doi.org/10.1101/2023.03.21.533623, 03.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459–482. https://doi.org/10.1002/cne.920180503