UNIVERSIDADE D
COIMBRA

Ana Rafaela Melanda Marques

Using the body to help to predict the brain: integration of body and brain physiological signals to decode visual attention and neural response to objects

Dissertation in the context of the Master's in Biomedical Engineering in the field of Bioinformatics, advised by PhD. Prof. Maria J. Ribeiro, PhD. Prof. Marco Simões and PhD. Prof. Davide Borra, presented to the Faculty of Sciences and Technology of University of Coimbra

September 2023

UNIVERSIDADE D
COIMBRA

University of Coimbra
Faculty of Sciences and Technology

Using the body to help predict the brain: integration of body and brain physiological signals to decode visual attention and neural response to objects

Ana Rafaela Melanda Marques

Dissertation in the context of the Master's in Biomedical Engineering in the field of Bioinformatics, advised by PhD. Prof. Maria J. Ribeiro, PhD. Prof. Marco Simões and PhD. Prof. Davide Borra, presented to the Faculty of Sciences and Technology of University of Coimbra

Coimbra, September 2023

Agradecimentos

Aproximando-me do término do meu percurso académico, não poderia de deixar de agradecer a todos os que, de forma direta e indireta, contribuíram para o meu crescimento, tanto a nível académico como pessoal.

Em primeiro lugar quero agradecer aos meus orientadores: à Professora Doutora Maria Ribeiro, ao Professor Doutor Marco Simões e ao Professor Doutor Davide Borra. Obrigada por me terem dado a oportunidade de desenvolver este projeto e, mais importante ainda, pela ajuda, orientação e paciência ao longo de todo o projeto.

Não posso deixar de agradecer ao CIBIT – Coimbra Institute for Biomedical Imaging and Translational Research, por me ter fornecido todos os recursos essenciais para a realização deste projeto. Quero também agradecer a todos aqueles que tiraram um bocadinho do seu tempo para serem voluntários neste projeto.

De seguida, agradeço à minha família pelo apoio incondicional que me deram ao longo da minha jornada académica. Em particular, quero agradecer aos meus pais, por me permitirem e incentivarem a perseguir os meus sonhos, e agradecer à minha irmã, por estar sempre lá para mim.

Não posso deixar de mencionar os meus amigos que, ao longo do meu percurso académico, foram uma fonte de apoio e tornaram estes cinco anos muito mais leves e divertidos. Um agradecimento especial àqueles que me apoiaram, encorajaram e tornaram este último ano menos desafiador: Ana Gabriela, Catarina e Carolina. Por fim, quero agradecer ao meu namorado pela paciência, pelo carinho e por todo o apoio durante este período.

Um obrigada a todos os que entraram na minha vida e me ajudaram a crescer em todos os níveis. Um bem-haja ao fim desta excelente etapa e ao início de outra que se avizinha.

Resumo

Introdução

A perceção de um determinado estímulo sensorial depende do nosso estado neural e fisiológico no momento do estímulo. Esse estado fisiológico pode ser alterado quando é induzido um estado de expectativa em antecipação de um evento relevante. Esse estado de expectativa está associado a uma ativação do córtex frontal, inibição motora, ativação do sistema nervoso parassimpático (desaceleração cardíaca), ativação do sistema nervoso simpático (dilatação da pupila e alterações na condutância da pele) e a uma facilitação da tomada de decisão percetual. No entanto, ainda não está claro se as alterações na fisiologia corporal afetam diretamente o processamento sensorial ou se são somente uma consequência das alterações cerebrais que por si afetam este processamento.

Objetivos

O objetivo deste estudo é investigar a interação entre a fisiologia corporal e a atividade cerebral e o seu impacto no processamento visual. Neste estudo usamos duas abordagens para investigar aspetos diferentes deste problema. Primeiro, investigamos como o estado de expectativa é refletido em mudanças nos sinais fisiológicos do corpo e do cérebro, e verificamos se esse estado pré-estímulo modula a perceção visual e de que maneira o os sinais corporais contribuem para essa perceção. Segundo, investigamos se o estado pré-estímulo modula a representação cortical de diferentes categorias visuais.

Métodos

Neste trabalho, utilizámos uma tarefa de discriminação visual, na qual os participantes tinham de determinar se o estímulo visual apresentado consistia num carro ou numa casa. Cada ensaio começava com um sinal sonoro de alerta que informava os participantes sobre a aproximação do estímulo visual. Enquanto os participantes realizavam a tarefa, foram adquiridos simultaneamente o eletroencefalograma (EEG) e sinais fisiológicos (eletrocardiograma, frequência respiratória, pupilograma e movimentos oculares). Para estudar o efeito do estado neural e fisiológico no processamento visual, usámos algoritmos de aprendizagem automática e testámos dois tipos de classificadores – máquina de vetores de suporte e redes neurais convolucionais.

Resultados

Os resultados sugerem que o estado de expectativa induz alterações no sinal de EEG e na fisiologia corporal. Em particular, o estado de expectativa caracterizou-se pelo aparecimento de um potencial negativo no sinal de EEG, por uma desaceleração cardíaca, por uma diminuição na ocorrência de pestanejos e movimentos oculares refletindo uma inibição motora, pela dilatação da pupila e pelo aumento da duração dos ciclos respiratórios. Para além disso, foi possível construir classificadores capazes de prever a deteção do estímulo visual usando como input o sinal de EEG pré-estímulo, a atividade respiratória ou o comportamento da pupila. Contudo, quando combinados com o sinal EEG, nenhum dos sinais fisiológicos forneceu informações suplementares ao classificador, o que sugere que os sinais fisiológicos não afetam a perceção visual diretamente. Por fim, avaliamos se o estado pré-estímulo modula a representação cortical das categorias visuais. Para isso construímos um classificador capaz de classificar a categoria visual tendo como input os potenciais evocados pelos estímulos visuais. No entanto, não encontrámos evidência de que esta classificação fosse modulada pela atividade cerebral ou corporal pré-estímulo.

Conclusões

Estes resultados sugerem que a atividade pré-estímulo influencia o reconhecimento de estímulos visuais. No entanto, não encontrámos evidência de que as alterações observadas na fisiologia corporal modulem a perceção visual ou a representação neural do estímulo.

Palavras-chave: atividade neural e corporal pré-estímulo; atenção; interações corpo-cérebro; Redes Neurais Convolucionais; Máquina de vetores de suporte

Abstract

Introduction

The perception of a sensory stimulus depends on the neural and physiological state at the time of the stimulus. This physiological state can be altered when a state of expectation is induced in anticipation of a relevant event. This state of expectation is associated with an activation of the frontal cortex, motor inhibition, activation of the parasympathetic nervous system (cardiac deceleration), activation of the sympathetic system (pupil dilation and changes in skin conductance) and a facilitation of perceptual decision making. However, it is still unclear whether changes in body physiology directly affect sensory processing or whether they are simply a consequence of neural changes that in turn affect this processing.

Goals

The goal of this study was to investigate the interaction between body physiology and brain activity and their impact on visual processing. In this study, we used two approaches to investigate different aspects of this problem. First, we investigated how the state of expectation is reflected in changes in body and neural physiological signals, and we verified whether this pre-stimulus state modulates visual perception and how body signals contribute to this perception. Second, we investigated whether the pre-stimulus state modulates the cortical representation of different visual categories.

Methods

In this work, we used a warned visual discrimination task where participants were required to determine if the presented visual stimulus consisted of a car or a house. Each trial started with an auditory warning cue that alerted participants of the upcoming stimulus. While participants were engaged in the task, we simultaneously acquired the electroencephalogram (EEG) and body physiological signals (electrocardiogram, respiratory rate, pupillogram and eye movements). To study the effect of the neural and body state on visual processing, we used machine learning algorithms and tested two types of classifiers – support vector machines and convolutional neural networks.

Results

The results suggested that the state of expectation induces changes in the EEG signal and in body physiology. In particular, the state of expectation was characterized by the appearance of a negative potential in the EEG signal, by cardiac deceleration, by a decrease in the occurrence of blinks and eye movements reflecting motor inhibition, by pupil dilation and by an increase in the duration of the respiratory cycles. Furthermore, it was possible to develop classifiers capable of predicting the detection of visual stimuli using as input the pre-stimulus EEG, respiratory activity or pupil behavior. However, when combined with the EEG signal, none of the physiological signals provided supplementary information to the classifier, which suggests that these physiological signals do not affect visual perception directly. Finally, we evaluated whether the pre-stimulus state modulates the cortical representation of visual categories. To do this, we developed a classifier capable of classifying the visual category using as input the potentials evoked by the visual stimuli. However, we found no evidence that this classification was modulated by pre-stimulus brain or body activity.

Conclusions

Overall, the results suggest that pre-stimulus activity influences the recognition of visual stimuli. However, we found no evidence that changes observed in body physiology modulate visual perception or the neural representation of the stimulus.

Key words: pre-stimulus neural and body activity; attention; body-brain interactions; Convolutional Neural Network (CNN); Support Vector Machine (SVM)

Table of contents

List of figures

List of tables

List of equations

Abbreviations

ACH Acetylcholine

AUC Area under ROC curve

BPM Beats per minute

CNN Convolutional neural network

CNV Contingent negative variation

ECG Electrocardiography

EEG Electroencephalography

ERP Event-related potential

ET Eye tracker

ICA Independent component analysis

i.e. That is

LC Locus coeruleus

MEG Magnetoencephalography

NE Norepinephrine

ROC Receiver Operating Characteristic

RBF Radial basis function

RT Reaction time

stP Single trial field potential

stBHA Single trial broadband high-frequency

SVM Support vector machine

1 Introduction

In this chapter, we will start with a contextualization and motivation of the thesis. Then, we will describe the theoretical background, as well as the research findings in the field of interaction between the brain and the body. Finally, in this chapter, a summary of the current study is made.

1.1 Contextualization and Motivation

The perception of a sensory stimulus depends on the neural and physiological state at the time of the stimulus. Numerous researchers posit that these neural and physiological states, observed by fluctuations in endogenous activity, have a significant influence on neural processing, i.e., these fluctuations can, for example, modulate stimulus processing (1), increasing or decreasing the chances one has to correctly process and categorize the stimulus content. Endogenous activity represents the activity that naturally occurs within an organism's body, that is, without the influence of an external stimulus.

Moreover, it is established that brain and body physiological activity such as cardiac activity, pupillary response (a measure of brain arousal), respiration and neural activity are associated with sensory processing, for example:

- Variations in heart rate have been correlated with changes in visual evoked responses, thereby suggesting a potential role of cardiovascular activity in the process of visual processing (2).
- Fluctuations in pupil size reflect changes in cognitive effort, with greater pupil dilation heightened cognitive effort and higher attentional level (3).
- Enhanced accuracy in visuospatial tasks is observed when stimuli are presented during inhalation in comparison with exhalation (4).
- Fluctuations in the internal state of the brain also affect sensory processing (5,6).

Attentive anticipation, the state of focused attention during sensory expectation, is associated with faster and more sensitive sensorimotor processing. This state is characterized by specific patterns of cortical activity, motor inhibition, cardiac deceleration and pupil dilation. In conclusion, this state of attention causes changes in brain and body physiology that are associated with both stimulus processing and behavioral performance (7–9).

Having stated this, several studies have already established an influence of fluctuations in physiological activity on visual perception. Nonetheless, most of these investigations have not tested whether the influence is content-specific or not. It was precisely this component that came to light in the studies conducted by Li, Y. et al. (1) and Podvalny, E. et al. (10), serving as the motivation for the present project. Considering this, both studies centered their efforts on studying the role of pre-stimulus activity in shaping the way visual stimuli are decoded (1,10).

- **Endogenous activity modulates stimulus and circuit-specific neural tuning and predicts perceptual behavior**

Li, Y. et al. (1), in *Endogenous activity modulates stimulus and circuit-specific neural tuning and predicts perceptual behavior,* studied the hypothesis that pre-stimulus activity modulates decoding accuracy in response to visual stimuli. To do so, they used implanted EEG electrodes to acquire neural activity data from 30 patients with epilepsy. During data collecting, participants performed a visual task. From the iEEG (intracranial electroencephalography) data, the authors extracted stP (single-trial field potential), stBHA (single trial broadband high-frequency) and the phases at different frequencies of the pre-stimulus activity. The stP represents the signal that was extracted from the data through a bandpass filter limited between 0.1 and 115 Hz. The stBHA activity was defined, on each trial, as the mean z-score of the power spectral density across 40–100 Hz. The power spectral density describes how the signal's power is distributed in frequency components (11). The task consisted of viewing images from different categories (faces, bodies, words, houses, hammers, and shuffled images) and pressing a button according to the stimulus category. Two hundred forty-six selective electrodes were used for each of the task categories, which allowed examining the effects of endogenous activity for each category.

The authors used a machine learning algorithm to discriminate the trial's category. With this analysis, the aim was to study how classification accuracy was influenced by pre-stimulus activity. To achieve this, they employed a regression logistic model to modulate classification boundaries based on the pre-stimulus activity. The algorithm used the pre-stimulus information to adjust classification boundary and, consequently, to optimize classification on each trial. The results demonstrated that the incorporation of pre-stimulus activity into the model enhanced the classification accuracy across all visual categories, outperforming the classification accuracy using only post-stimulus activity. These results suggest that pre-stimulus neural activity modulates the stimulus representation (1).

Having established the influence of pre-stimulus activity on visual processing, the authors proceeded to investigate the underlying mechanisms of this influence. From the

regression logistic model, the authors extracted a trial-specific measure of how much influence pre-stimulus activity has on neural decoding accuracy. This measure was called "modulation index" and was extracted for each electrode. By correlating the modulation index values across pairs of electrodes, the authors found a weak correlation between electrodes. Nevertheless, for electrodes that share the same category-selectivity this correlation was statistically significant, whereas for electrodes of different category-selectivity it was not statistically significant. Furthermore, comparing the results to pairs of electrodes, they also verified that the neural state modulates, in a specific manner, the post-stimulus activity, particularly within regions specialized for the processing of the visual stimulus under consideration (1).

To test the hypothesis that the same aspect of the pre-stimulus activity that influences decoding accuracy correlates with behavioral performance, the authors studied the correlation between trial-specific modulation index and behavioral reaction time. The pre-stimulus modulation index and reaction time (RT) were found to be statistically correlated; however, this correlation was not verified between RT and the post-stimulus features. This finding suggests that pre-stimulus neural activity has a greater influence on reaction time than post-stimulus neural activity (1).

Finally, the authors evaluate the contribution of different aspects of the pre-stimulus features. Their research revealed that when a stimulus is presented during a period of relatively low endogenous activity within regions selective to that type of stimulus, this could be an indication of reduced pre-stimulus noise. This low endogenous activity was evidenced by lower pre-stimulus averages and variances of stBHA and stP (1).

- **A dual role of prestimulus spontaneous neural activity in visual object recognition**

In *A dual role of prestimulus spontaneous neural activity in visual object recognition*, the authors also studied the influence of pre-stimulus neural state (12). In this paper, the authors hypothesized that the recognition of a given object could be influenced by two main models: the general model and the specific model. In the general model, the authors hypothesized that pre-stimulus brain states exerted an influence on the recognition of the stimulus regardless of the stimulus content. The specific model, in turn, suggested that pre-stimulus brain states enhance the process of recognition in a category-specific manner. Data were acquired on 25 participants with magnetoencephalography while performing a visual task. Magnetoencephalography is a functional neuroimaging technique that records the magnetic fields created by electrical currents that occur naturally in the brain (12). The task consisted of presenting real or "scrambled" images, with the real images having a low contrast

to achieve a 50% subjective recognition rate. After stimulus onset, the participants had to report the category of the object and their recognition of the stimulus. It should be noted that they were asked to predict stimulus category even if they did not see any object. The object categories were faces, animals, houses, and man-made objects (10).

To validate the suggested models, the authors used a logistic regression model, which received input vectors containing pre-stimulus activity, the categories of objects and subjective recognition reports. The pre-stimulus activity was defined as the average of the magnetoencephalography signal over a two-second interval before stimulus presentation in each trial. The authors started by studying the influence of pre-stimulus activity in visual recognition according to the general model. Their analyses revealed that the pre-stimulus neural state influences the recognition of an upcoming stimulus regardless of category (recognition of the presence of an object). In other words, the general model proposes the existence of a general neural mechanism that is not specific to any particular stimulus category but still exerts an influence on subsequent recognition. Remarkably, the researchers found that the subjective recognition rate increased 7.2%, when the participant reported recognizing an object in a previous trial. This observation suggests that the detection of a stimulus results in fluctuations in neural activity that facilitate the recognition of the subsequent stimulus (10).

Furthermore, the authors conducted a specific model to ascertain how pre-stimulus activity enhances stimulus recognition in a specific way based on the category of the stimulus. Their findings suggest that a particular brain state that facilitates the recognition of a stimulus belonging to category “a” is different from the state that facilitates the recognition of a stimulus from category "b”. Curiously, they discovered that recognition of the stimulus becomes easier when the brain correctly predicts a specific stimulus category (and the expected category aligns with the actual category of the stimulus). Summarizing, they conclude that the neural activity carries predictive information concerning the content of the forthcoming recognized stimuli, consequently increasing the sensitivity of predictive mechanisms (10).

Finally, the authors also related the two processes, the general and specific model, with the pupil size. This decision was motivated by the fact that pupil size serves as a measure of arousal fluctuations. They verified that, in a general way, pre-stimulus neural process correlates with pupil-linked arousal. However, this correlation is not specific to the category (10).

In conclusion, the authors of both studies verified the existence of pre-stimulus activity fluctuations that modulate stimulus recognition, with distinct fluctuations favoring the recognition of different categories. Using these discoveries as a motto, we decided, in addition to studying neural activity, to also incorporate body physiological activity. It is still unclear whether changes in body physiology directly affect sensory processing or whether they are simply a consequence of neural changes that in turn affect this processing.

That said, the goal of this study was to investigate the interaction between body physiology and brain activity and their impact on visual processing. In this study, we used two approaches to investigate different aspects of this problem. First, we investigated how the state of expectation is reflected in changes in body and neural physiological signals, and we verified whether this pre-stimulus state modulates visual perception and how body signals contribute to this perception. Second, we investigated whether the pre-stimulus state modulates the cortical representation of different visual categories.

1.2 Visual perception

The individual interpretation and awareness of information obtained through our primary senses is referred to as perception. Perception refers to the processing of a specific form of energy that provides information about the environment, and it is influenced by prior knowledge, memory, and previous experiences (13).

Visual perception is the ability to perceive the world around us through the light that reaches our eyes. We can describe visual perception as the process where a retinal image is transformed into a representation of the external world (14). It is the visual system that, through specialized cortical areas, is responsible for visual perception (13).

1.2.1 Visual system

The visual system encompasses the eye as a sensory organ, as well as the visual cortex and the optic nerve, parts of the central nervous system. The light energy enters the eye through the cornea, then passes through the pupil and lens, responsible for focusing the light onto the retina. The retina is composed of specialized cells responsible for visual transduction: the rods, which are sensitive to dim light, and cones, specialized in transducing bright light. The activation of photoreceptors occurs when light reflects on the sclera. The photons are then observed by the photoreceptors, causing their excitation. In turn, these cells connect to bipolar cells, which, when activated, induce action potentials in retinal ganglion

cells. The axons of these ganglion cells emerge from the retina to create the optic nerve, facilitating the transmission of electrical impulses from the eyes to the brain (15,16).

The optic tract then diverges into two visual pathways: the geniculostriate pathway and the tectopulvinar pathway, both of which relay through the thalamus. The principal pathway, the geniculostriate pathway, is involved in processes such as visual consciousness and pattern recognition. This pathway connects the retina to the lateral geniculate nucleus, which connects to the visual cortex. On the other hand, the tectopulvinar pathway plays an important role in detecting and orienting visual attention. The tectopulvinar pathway relays from the eye to the superior colliculus in the midbrain tectum and, subsequently, reaches the visual areas in the temporal and parietal lobes through relays in the lateral posterior-pulvinar complex of the thalamus (15,16).

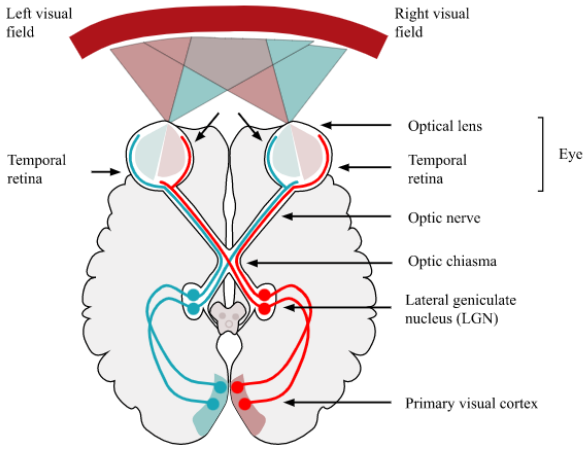


Figure 1.1 – *Visual pathway. Adapted from* (17).

1.2.1.1 Visual cortex

The visual cortex, situated within the occipital lobe of the cerebral cortex, is the first cortical region responsible for processing visual information. The sensory information reaches the visual cortex after passing through the lateral geniculate nucleus of the thalamus. Notably, each hemisphere of the brain has its own visual cortex. The left hemisphere, for instance, receives information from the right visual field and vice versa (18,19).

The visual cortex is responsible for receiving, dividing, and integrating visual data. Then, the processed information is delivered to other regions of the brain (18,19).

Five distinct areas compose the visual cortex that can be divided into two large groups based on their function and structure:

- Primary visual cortex (V1) – area that receives the sensory input from the lateral geniculate nucleus.
- Extrastriate cortex – composed of V2, V3, V4 and V5.

It is believed that, throughout the transmission of visual information, each cortical area is increasingly specialized, and as a result, contains more specialized cells. These brain regions learn to react to specific objects, facilitating rapid recognition of previously seen objects (18).

The V1 region of the visual cortex, which has six unique layers, is the region responsible for receiving and processing visual information. Each layer is characterized by a different cell type and function. Basic visual features, such as position and direction, are encoded in the responses of V1 neurons (18). This region, more properly the fourth layer, receives information that flows from the optic nerve to the lateral geniculate nucleus and then to this area (20). Neurons within the V1 area are organized into columns across the cortex's thickness. Each column contains neurons with similar preferred orientation, suggesting that neurons in one column might react to visual stimuli with a specific type of orientation, while neurons in another column might react to a different orientation, for example (18,19).

Subsequently, V2 receives feedforward signals from V1 and sends feedback signals to the V1 area. By integrating the input it receives, this V2 is able to code higher levels of visual complexity. Cells in this region are believed to respond to characteristics such as spatial frequency, object orientation and differences in color (18,19,21). The data, after being processed and integrated in this area, is sent to more anterior regions by two different pathways known as the dorsal and the ventral streams. Both pathways are specialized in processing different components of visual data (18,19,21). The ventral stream, often linked to object recognition, goes through the V3 and V4 areas before culminating in the Inferior Temporal Cortex. On the other hand, the dorsal stream focuses on spatial processing and visual-motor skills. It travels through V3 and V5 areas reaching the Posterior Parietal Cortex (18,19,21).

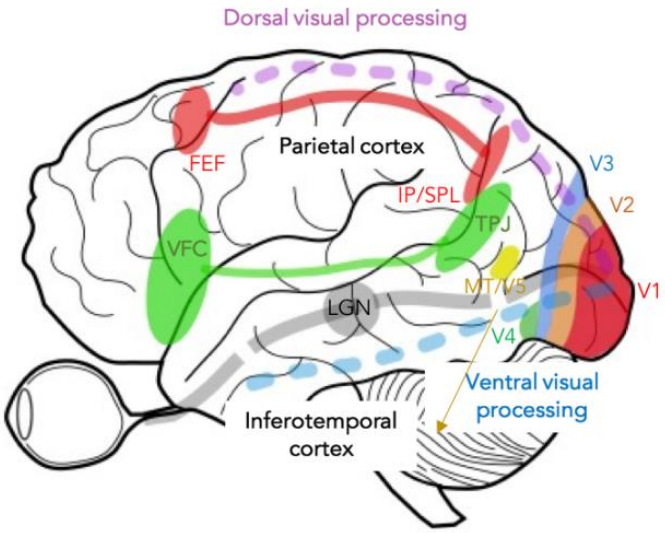


Figure 1.2 *- Visual processing pathways (LGN - lateral geniculate nucleus; VFC - ventral frontal cortex; IP/SPL - intraparietal/ superior parietal lobule; FEF - frontal eye fields; TPJ - temporoparietal junction). Adapted from* (22).

1.2.2 Visual processing

Visual processing is the ability of the brain to use and understand visual information derived from the environment. This highly specialized process enables the brain to recognize objects and patterns with no conscious effort. Numerous brain areas and higher-level processes work together to transform light energy into a meaningful image (18,23).

Visual processing can be divided into three levels: low-level, mid-level and high-level vision. Low-level visual processing is thought to involve the representation and analysis of basic features, such as local color, luminance, or contrast (18,23). Then, we have the mid-level, which refers to the representation of the interactions between basic characteristics and properties, including surfaces (textures), higher-order image statistics, disparity, and intermediate form features (18,23). Finally, high-level visual processing refers to the cognitive mechanism responsible for integrating information from various sources into the visual information that is represented in the consciousness of each individual. It includes complex tasks like object processing, which necessitate the integration of diverse visual information (18,23).

1.2.3 Modulation of visual processing

The perception that we have of a stimulus is, in addition to what was mentioned above, influenced by individual expectations and attention mechanisms. It is believed that visual perception results from the interaction between two types of processing: bottom-up and top-down processing. Bottom-up processing refers to the visual system's ability to utilize the incoming visual data, through the automatic processing of basic sensory information (14,24). In turn, top-down processing refers to more complex cognitive functions influenced by our prior knowledge of the world, as well as our expectations and goals (14,24). As a result, top-down processing changes the sensitivity of neurons to specific stimuli, which affects the information transmitted by them. The information carried by the top-down signal facilitates the interpretation of the visual world, playing a role in the encoding and recall of learned information (14,24).

Initially, top-down processing was thought to be a modulatory mechanism that reflected how much attention was given to different things. However, it is now believed that top-down processing interacts with bottom-up information to "optimize" processing by incorporating higher-level object representations, thereby exerting a significant impact on perception (25).

1.3 Physiological signals and perception

We are constantly exposed to multisensory stimuli coming from different sensory sources. These sources can be external, such as visual, auditory or tactile signals, or internal, also known as interoceptive signals, such as cardiac, respiratory and visceral inputs. The relationship between body and brain signals and visual perception is complex and bidirectional, with both mutually influencing and molding each other in diverse ways (26).

It is hypothesized that our neural and physiological state at the moment of the stimulus onset influences our recognition and perception of it. Moreover, previous research has shown that changes in brain and body physiology are associated with changes in the speed and sensitivity of cognitive reactions to external stimuli.

1.3.1 Cardiac activity

1.3.1.1 Cardiac activity and electrocardiography

The cardiac cycle is responsible for maintaining blood flow throughout the body by rhythmically pumping blood. To do this, the heart muscle must contract and relax along with the opening and closing of the heart valves (27,28). The cycle is composed of two phases: the contraction phase, known as systole, and the relaxation phase, known as diastole (27,28). This cardiac activity results from a sequence of complex electrophysiological events, which can be measured in the electrocardiogram (ECG) (27,28).

The ECG consists of measuring the electrical activity of the heart generated by the cardiac cycle. The heartbeats produce action potentials capable of being detected through electrodes (29). The ECG signal is composed of a distinctive waveform corresponding to different phases of the cardiac cycle (Figure 1.3). The P wave represents atrial electrical activity, that is, it represents the propagation of depolarization through the atria. Subsequently, the QRS complex appears, reflecting the electrical depolarization of the ventricles, that is, the moment of blood ejection into the arteries. Following the depolarization of the ventricular cells, their repolarization occurs, that is, the ventricular fibers begin to relax, which represents the T wave. It is important to note that there is no specific wave that represents the repolarization of the atrial cells, since this process occurs simultaneously with the ventricular depolarization and is not large enough to generate a wave on the electrocardiogram (27–29).

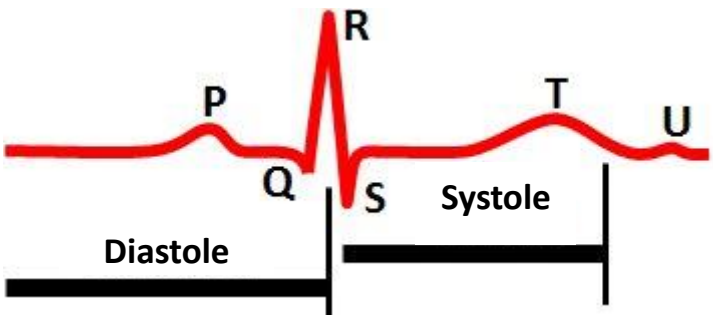


Figure 1.3 *– The electrocardiogram – example of ECG waves. Adapted from* (30).

Heart rate is the speed of the cardiac cycle, and it is measured by the number of contractions per minute. It can vary with the different physiological requirements of the organism. The regulation of heart rate can be done through two distinct mechanisms: intrinsic control or extrinsic control. In intrinsic control, the heart rate is regulated by the heart itself in response to changes in the volume of blood passing through the heart. The extrinsic control, in turn, is the result of the action of the autonomic nervous system (28). The

sympathetic division of the autonomic nervous system can induce an increase in heart rate, as well as an increase in the strength of the heart's contraction; on the contrary, the parasympathetic system leads to a decrease in cardiac activity (28,31).

The neural mechanisms that control cardiac activity are already known. However, the effect of changes in cardiac activity, such as the deceleration observed during freezing states, on sensory processing remains unknown.

1.3.1.2 Impact of cardiac activity on perception

Several authors demonstrated the impact of presenting a stimulus at specific moments of the cardiac cycle on the way the stimulus is processed. In *Insula Mediates Access to Awareness of Visual Stimuli Presented Synchronously to the Heartbeat* (32), for example, the authors observed that when a visual stimulus is presented synchronously with the cardiac cycle, stimulus discrimination is less precise, and it requires more time for the individual to perceive the presentation of the stimulus.

Researchers have found that a subject is more likely to miss the visual stimulus presentation when it is presented during the systolic phase than when it is presented during diastole (32). Various explanations have been proposed to understand this relationship. One notable finding is the relationship between the perception of a given stimulus and the activity of baroreceptors, which are pressure sensors. This activity is influenced by the cardiac cycle, with greater activation of baroreceptors during systole, for example. It has been widely believed that the activation of baroreceptors exerts a general inhibitory effect on the central nervous system, leading to a suppression of fundamental sensory and sensorimotor processes. Thus, during systole, there is an increase in baroreceptor pressure, resulting in the inhibition of the central nervous system (33).

Additionally, according to Sandman et al. (2), lower heart rates may facilitate the perception of visual stimuli. They found that variations in heart rate are related to changes in visual evoked responses, more precisely at the level of cerebrovascular and electrocortical activity. The authors verified that cephalic pulse amplitudes were largest when heart rate was low and smallest when heart rate was high. These findings suggest a potential role of cardiovascular activity in the process of visual processing.

In turn, more recent experiments did not observe any relationship between visual processing and the cardiac cycle (34).

1.3.2 Respiratory activity

1.3.2.1 Respiration and its evaluation

Respiration is the process responsible for delivering oxygen to tissues and removing carbon dioxide. For this, it is necessary the movement of air into and out of the lungs, promoting gas exchange between the inhaled air and the circulatory system. This process is called pulmonary ventilation (28). The pulmonary ventilation process consists of two distinct phases: inspiration, involving the intake of air, and expiration, that consists in the expulsion of air. These movements are controlled by the activity of the diaphragm, which contracts and expands in coordination with the intercostal muscles, allowing the entry and exit of air from the lungs (28,29).

The respiratory rate refers to the complete respiratory cycles occurring within a defined timeframe, typically one minute. Adults typically breathe between 12 to 20 breaths per minute (28).

Over the years, several methods have been developed to investigate pulmonary ventilation. To measure pulmonary ventilation, we can analyze the air that enters and leaves the lungs through spirometry (28,29,35). Another way to study pulmonary ventilation is through the displacement of the chest caused by the activity of the diaphragm (28,29,35). This non-invasive technique involves the placement of sensors on the chest or abdomen to detect changes in the position or displacement in these structures during the entry and exit of air from the lungs. In our project, we used a common and non-invasive technique known as respiratory inductance plethysmography. This technique consists of measuring, with recording bands around the thorax and abdomen, changes in circumference as the respiratory muscles contract and relax (28,29,35). These changes in circumference are then used to calculate changes in lung volume, allowing us to assess pulmonary ventilation without invasive procedures.

1.3.2.2 Impact of respiration on perception

Perl, O. et al. (4) demonstrated that the acquisition of visuospatial stimuli synchronized with inhalation optimizes stimulus processing. Inhalation was found to affect the activity of the primary olfactory cortex, which, in turn, impacts cognition that depends on limbic structures. It is believed that fluctuations in these limbic structures regulate cortical excitability and coordinate network interactions (36). The authors further concluded that individuals have the capacity to modulate their breathing to match the stimulus presentation, which improves the processing of sensory information (4).

1.3.3 Pupillary response

1.3.3.1 Pupil and eye tracker

Eyes are organs of the visual system capable of providing living organisms the ability to perceive and process visual information. The human eye has a spherical shape and comprises several essential structures, including the cornea, sclera, pupil, iris, lens, retina, and optic nerve (37). The pupil, located at the center of the iris, is responsible for allowing and regulating the light that reaches the eye. To control the influx of light, the iris adjusts the pupil's diameter and, consequently, the pupil's size (37).

The size of the pupil can be modulated through changes in luminosity. To this end, light-sensitive retinal cells in the ocular system react to bright light by sending signals to the oculomotor nerve. Rod and cone photoreceptors, as well as melanopsin ganglion cells are examples of light-sensitive retinal cells (38). These signals are sent to the parasympathetic branch of the oculomotor nerve, which terminates on the circular iris sphincter muscle. It is the contraction of this muscle that will produce a reduction in pupil size. In turn, pupil dilation is moderated by the sympathetic nervous system, which controls the iris dilator muscle (38). This process is known as the pupillary light reflex. Furthermore, pupil dilation can also occur in response to stimuli that evoke interest or arousal (38).

Pupil size variations under conditions of constant luminance can offer information about cognitive and emotional processes. We can measure the pupil size and reactivity through pupillometry, a non-invasive technique that measures the dynamic changes in pupil size over time (39). This technique can be performed using an infrared video camera to record pupil's behavior. As light enters the eye, it reflects off the retina and exits the eye, returning to the camera (40). To record the pupil behavior, we can use an eye tracker (ET).

Eye tracking is a technology that involves monitoring eye movements in relation to the head or the point of gaze (where someone is looking) (3,41). It is commonly used to study visual perception, cognitive processes and attention. There are several techniques to measure eye movements. The most popular method extracts the eye position from video images. Eye trackers are often affordable and provide enough temporal resolution and precision to detect even little fluctuations in pupil size (3,41).

1.3.3.2 Impact of modulation in pupil-linked arousal on perception

The pupil reacts to both light and internal mental states. For instance, excitement or bad experiences that modulate arousal state might cause pupil dilation. Other common arousal-related processes that influence pupil size include exploratory behavior during difficult tasks and decision-making processes (42).

According to published research, increasing task demands result in larger pupil size. So, changes in pupil size can reflect differences in cognitive effort, with greater pupil dilation indicating higher cognitive effort and attentional focus (3). Some studies suggest that, for example, in cognitive control tasks, pupil dilation can help to predict task performance, that is, participants who have higher pupil dilations perform better on these tasks than those who have smaller dilations (3,43).

1.3.4 Blinking activity

1.3.4.1 Blinking and eye tracker

With an ET, we can also measure blinking. Blinking is a semi-autonomic rapid closure of the eyelids that helps to protect the eyes by cleaning and wetting them (44). When blinking occurs, the tear glands secrete salty fluids that coat the eyes, facilitating the removal of tiny dust particles and maintaining the exposed part of the eyeball wet. Under normal circumstances, the average person blinks between 12 and 15 times per minute. However, there are circumstances, such as being in a room filled with smoke, when we blink more frequently to keep our eyes wet and clear (44). Furthermore, recent studies propose that blinking allows us to shift our focus from one task to another, "resetting" our attention mechanism. Therefore, it is believed that blinking might be associated with cognitive processes (45).

1.3.4.2 The relationship between blinking and sensory perception

Eye blinking serves not only to lubricate the cornea, but also as a behavioral manifestation of central nervous system activity. Studies have demonstrated that blinking can affect task performance, particularly by interfering with attentional focus. Blinking causes individuals to momentarily close their eyes, interrupting the flow of visual information, and lose focus (46). In this way, while performing a task, it is observed a reduction in the frequency of blinking to avoid affecting visual detection and discrimination (47).

It is believed that the blink rate may be an indicator of cognitive load and task engagement. Subjects tend to blink less in demanding cognitive tasks, which suggests a higher attentional focus and cognitive effort. On the other hand, the blinking rate tends to increase when the task demand is lower or when subjects experience mental fatigue. This modulation of blink rate suggests a decrease in cognitive effort and attentional focus (48,49).

1.3.5 Saccadic activity

1.3.5.1 Saccades and eye tracker

Saccades are responsible for rapidly shifting the point of fixation from one object to another, through rapid and jerky eye movements. They occur automatically whenever the eyes are open, even when fixed on an object. Despite being an automatic mechanism, saccades can be intentionally induced. The amplitude of these movements can range from slight motions made while reading to more significant movements made while gazing around a room (50).

Saccades are rapid eye movements toward visual, auditory, or tactile stimuli. Their purpose is to direct the focus onto objects of interest and bring them into clear view on the fovea, the central area of the retina providing the highest visual acuity (51). However, vision is impaired during saccades due to two reasons. Firstly, during large saccades, the image moves very quickly causing blurriness, which can impair vision. Secondly, at the initial part of each saccade, there is a vision obstruction, a blanking-off process called saccadic suppression. In between saccades the eyes remain still during fixations (52). As in the pupil section, to measure saccadic activity we also can use an eye tracker.

1.3.5.2 The impact of saccades on perception

Saccades are movements of the fovea between points of interest that help us to build a comprehensive perception of our surroundings. This enables us to efficiently and quickly locate crucial information (53). However, saccades, in certain tasks, can be a source of errors, impairing vision. For instance, during saccades there is a rapid motion of the image, which may result in blurriness and loss of crucial information. Therefore, if a stimulus is presented during a saccade, there is a possibility that the individual may miss significant information related to the stimulus (47,52).

Studies conducted by Abeles D. et al. (47) have demonstrated that during attentive anticipation of sensory stimuli there is a reduction in saccade rates and that when saccades were performed during target presentation, accuracy rates were lower, even for auditory stimuli suggesting a general impact on sensory perception.

1.3.6 Neural activity

1.3.6.1 Neural and electroencephalography

The nervous system is the major system for control, regulation, and communication in the body. It is essential to all mental processes, such as memory, learning, and thinking (15,28,54). In conjunction with the endocrine system, the nervous system regulates and maintains homeostasis. Through millions of sensory receptors, this system gives us the ability to stay connected and responsive to both our internal and external environments (15,28,54). These receptors are sensitive to both internal and external changes, allowing the monitoring of a lot of information, such as temperature, light, pH levels or even carbon dioxide concentrations. The collection of all this information is called sensory input. The sensory input is converted into electrical signals that are transmitted to the brain. In the brain, different regions communicate and collaborate to perform complex cognitive tasks. This process is called integration. Finally, the nervous system sends impulses to different parts of the body to produce responses to the sensory input (15,28,54).

In the nervous system, the information flows from one neuron to another via an electrical potential known as an action potential. This action potential propagates along the axon to the synapse. Within the synapse, a small gap between neurons, neurotransmitters are released into the synaptic cleft, where they bind to receptors on the receiving neuron. Consequently, we observe two types of potentials: the action potential that flows through the neuron and the postsynaptic potential generated at the synapse (15,28,54). Action potentials have a very small extracellular amplitude, which makes their detection difficult. To produce an EEG signal (voltage differences captured outside the head by electrodes placed in the scalp), it is necessary to have a large number of simultaneous action potentials. Unlike action potentials, which last only one millisecond, postsynaptic potentials last about ten milliseconds. This increase in duration allows for potential changes to accumulate and to be registered extracellularly from the scalp (55).

EEG is a non-invasive technique used to measure the electrical activity of the brain. Through electrodes placed on the scalp, this technique measures the voltage fluctuations resulting from the postsynaptic potentials (56,57). There are two main types of EEG: monopolar and dipolar. In the monopolar EEG, the voltage difference is determined between an active electrode and a reference electrode. The dipolar EEG, on the other hand, corresponds to the voltage difference between two electrodes, both located on the scalp (58).

The electrodes are placed in standard positions, according to the International Federation of Clinical Neurophysiology (Figure 1.4). For this, specific anatomical points of the skull that are stable are used, such as the "nasion" (placed between the forehead and the nose), the inion (located at the back of the skull) and the preauricular point (58).

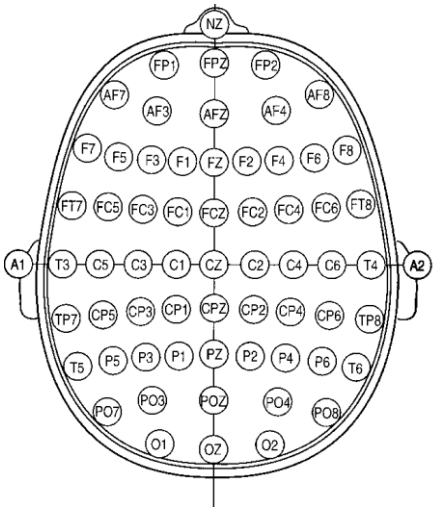


Figure 1.4 *– Electrodes positions in the International Federation of Clinical Neurophysiology system. Adapted from* (59).

The EEG wave can have varied forms, and can be classified according to its frequency, amplitude, shape, or even the position of the electrodes. In terms of frequency, they can be classified as:

- Delta waves (0.1 – 4 Hz) - predominantly associated with deep sleep, also called slow-wave sleep. During this stage of sleep, these become the dominant brain wave pattern (57).
- Theta waves (4 – 8 Hz) – characterized by their relationship with subconscious activity. These brain waves are commonly associated with a state of deep relaxation, drowsiness, and meditative states (57).
- Alpha waves (8-13 Hz) - typically associated with a state of relaxed wakefulness and are often observed in individuals who are engaged in activities that do not require intense mental effort (57).
- Beta waves (13-30Hz) - associated with various behaviors and actions. This type of waves occurs during states of consciousness such as talking, solving problems, judging, making decisions, among others (57).
- Gamma waves (30 – 100Hz) – associated to processes related to perception and consciousness. They tend to manifest during a state of hyper-alertness, facilitating the integration of sensory inputs (57).

1.3.6.2 The impact of neural activity on perception

Multiple authors demonstrated that the same stimuli may provoke different perceptions. They postulate that this perceptual variability is influenced by fluctuations in

neural activity (5,6). For example, alpha oscillations are thought to play an important role in sensory processing and perception by modulating neural excitability through functional inhibitory mechanisms (60,61).

In (61), the authors demonstrated how alpha oscillations affect the process of visual target detection. It was shown this influence was different from the influence exerted by endogenous factors, such as attention. The authors found a significant difference in the mean alpha power between trials where the target was not successfully identified and trials where the target was accurately detected.

These findings are consistent with those presented in (60). Using a visual detection task, the authors verified that reduced pre-stimulus alpha power, in early visual areas, was associated with increased perceptual sensitivity and enhanced discernible information from neural activity patterns. Moreover, it was found that the phase of oscillatory alpha activity immediately before stimulus presentation had an impact on trials' performance.

Summarizing, these findings suggest that fluctuations in endogenous neuronal activity affect sensory processing.

1.4 Alertness state

Attention can be described as the ability to select behaviorally relevant stimuli, responses, memories, or thoughts from among the many others that are irrelevant. According to Posner's paradigm, the attention system involves three major roles: orienting to sensory stimuli, which directs attention towards specific locations; executive function, which pertains top-down conflict detection and the suppression of distracting information; and maintaining a state of alertness (62).

To study the state of alertness, it is necessary to consider three distinct concepts: arousal, vigilance, and alertness. Alertness refers to a state of wakefulness and responsiveness to stimuli, whereas vigilance denotes the ability to maintain attention and sustain a high level of cognitive performance (63). Arousal, on the other hand, reflects the degree of cerebral cortex activation and can be associated with both alertness and vigilance (63).

The alerting network plays a crucial role in both the maintenance and preparation of attention, allowing individuals to focus on specific and predefined stimuli. So, in alertness, individuals are in a state of preparedness that enables them to react to stimuli. The level of

alertness significantly impacts our daily activities, such as learning, solving problems and remembering information (64).

Alertness can be classified into tonic and phasic alertness. Tonic alertness, mediated by the right frontal-parietal-thalamic network, involves the intrinsic regulation of arousal without external stimuli, i.e., it refers to the maintenance of an alert state. Research on circadian rhythms has demonstrated fluctuations in the tonic alertness throughout the day. For example, during daytime, there is an increase in the levels of tonic alertness, and then they fall during the evening and early morning hours (62,64,65). On the other hand, phasic alertness, which is mediated by the thalamic-mesencephalic regions (63), is characterized by the ability to temporarily enhance response readiness after an external stimulus, through the activation of the cognitive system (65).

In sensorimotor tasks, the existence of a warning signal before the target stimulus causes an increase in levels of alertness, increasing preparedness and improving the perceptual decision making (66). Additionally, some researchers have found that this induced state of alertness causes a reduction in reaction time (65).

- **Contingent negative variation**

The phasic changes in alertness are characterized by a negative shift in the EEG activity, which reflects cognitive processes related to the anticipation of an upcoming event (67,68). This slow negative shift, known as Contingent Negative Variation (CNV), was first identified by Walter et al. (67) and it appears to arise in the anterior cingulate cortex and adjacent structures (69).

These Event-Related Potentials (ERPs) occur between two paired stimuli: S1 and S2, which represent the initial and subsequent stimuli, respectively. The initial stimulus serves as a warning or preparatory cue, while the subsequent stimulus requires the subject to perform a response. In this way, it is believed that the CNV exhibits at least two associative functions: attentive orienting to the warning cue, induced by S1, and anticipatory attention during executive control, which is caused by S2 (68). These functions align with the two main psychophysiological components of the CNV waveform: the early CNV, which is thought to reflect the orienting response, and the late CNV, believed to indicate an anticipatory activity for the upcoming stimulus and preparation for motor response (Figure 1.5) (69).

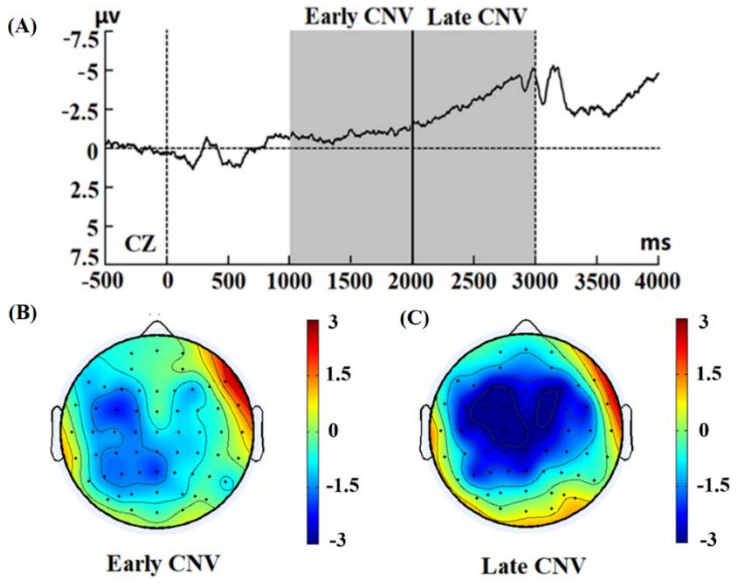


Figure 1.5 *- Event-Related Potential. (A) Grand average ERP locked to cue stimuli at the CZ electrode. (B) Topographic scalp distribution of early CNV. (C) Topographic scalp distribution of late CNV. Adapted from* (70).

It is believed that the alertness network involves fronto-parietal regions, mostly in the right hemisphere, and also some brainstem areas such as the locus coeruleus (LC). Studies using brain lesions have demonstrated that lesions in the right hemisphere had a greater impact on vigilance task performance than lesions in the left hemisphere. These findings suggest that the mechanisms necessary for maintaining the alert state seem to be present in the right hemisphere (62,65,71,72).

1.4.1 Anatomical pathways

It is believed that there are two principal regions that are involved in alertness: the anterior cingulate cortex and the dorsolateral prefrontal cortex. The dorsolateral prefrontal cortex is involved in detecting salient events, preparing and inhibiting motor responses, shifting and sustaining attentional focus, and maintaining alertness. The anterior cingulate cortex region, in turn, appears to mediate executive and attentional activities, conflict monitoring, and the capacity to activate, facilitate, and maintain reactions. According to a few neuroimaging studies, the noradrenergic system regulates these regions. The noradrenergic system originates in the LC structure and projects, principally, to frontal areas. Posner and Petersen (62) proposed that the parietal cortex, beyond the frontal areas, plays a crucial role in alertness (62,65,71,72).

In a PET investigation concerning an intrinsic alertness task, Sturm, W. et al. (71) reported significant activation in the right frontal cortices (particularly, the anterior cingulate cortex and the dorsolateral prefrontal cortex), as well as the inferior parietal cortex, thalamus, and brainstem (65,71). There is a connection between the dorsolateral prefrontal cortex and the inferior parietal cortex that might be completed by two other indirect pathways (via the anterior cingulate cortex and via the thalamus). Understanding these anatomical pathways is crucial for the development of a comprehensive network model of alertness and for understanding how different brain regions interact during alertness tasks (65,71).

1.4.2 Neurotransmitter systems involved in the modulation of alertness

All systems involved in the state of alertness, namely the neurotransmitter systems, exhibit extensive interconnections with direct projections between numerous brain areas (63). Different behavioral states are produced by variations in this network's activity. During drowsiness, slow-wave sleep, and surgical anesthesia, cortical networks exhibit a state called "synchronized" or "inactivated", characterized by slow, large-amplitude oscillations that demonstrate substantial synchronization between neuronal populations (73). On the contrary, during arousal, alertness, and paradoxical sleep, we have the "desynchronized" or "activated" state, where cortical networks are characterized by the absence of synchronous slow oscillations (73).

Cortical activation is induced due to the cortical action of neuromodulators, such as norepinephrine (NE) and acetylcholine (Ach) (73).

1.4.2.1 The role of the Locus Coeruleus-norepinephrine system in the modulation of alertness

There is a belief that the noradrenergic system originating from the LC structure may have a significant role in the maintenance of alertness. This structure, a modulatory nucleus located in the brainstem, has projections to the entire brain, and it is responsible for controlling the majority of NE release. NE represents the principal chemical component involved in the regulation of attention (62,74).

The LC plays a critical role in regulating arousal and attention. It has been observed that activation of the LC system, such as in response to a stimulus, for example, causes an

increase in NE levels in the brain, which results in a general increase in cortical activity and, consequently, enhanced alertness (75). In this way, it can be deduced that higher levels of NE contribute to higher levels of alertness.

The LC-NE system operates in two modes: tonic (baseline) and phasic (responsive to stimuli). When the tonic activity of the LC increases, the individual experiences enhanced attention and alertness, leading to improved task performance (63,75). The enhancement of sensory processing is believed to emerge from two mechanisms: an increase in the signal-to-noise ratio in neurons and a change in the receptor field properties, which decreases the threshold for response. However, if the increase in LC activity exceeds a particular threshold, individuals may become hyperactive and enter in a more easily distractible state (63,75).

Allied to this, increased levels of NE can also trigger activation of the autonomic nervous system, through direct projections to the spinal cord and projections to autonomic nuclei. In general, the LC increases sympathetic activity via the activation of α_1-adrenoceptors on preganglionic sympathetic neurons (63,75). In this way, the LC plays a major role in regulating the sympathetic nervous system and the physiological responses associated with this system, such as a reduction in salivary glands activity, an increase in heart rate and blood pressure, and pupil dilation (63,75). It is important to emphasize that the regulation of the sympathetic nervous system is a complex process, involving the interaction of different brain regions (63,75).

1.4.2.2 The role of the Cholinergic systems in the regulation of arousal and attention

Acetylcholine (ACh) is a neurotransmitter that plays a crucial role in the cholinergic system. Previous investigations employing lesion and microdialysis techniques supported the hypothesis that cortical cholinergic projections are imperative for performance in tasks involving attentional functions (63,76). Lesions in the basal forebrain (consist of several structures responsible to produce ACh) impair performance in tasks in which a state of attention is required, which suggest that the cholinergic system has an important role in the modulation of cognitive function, namely attention (63,76).

As in the previous section, we can also divide the cholinergic activity in tonic and phasic activity. The first one refers to the continuous release of ACh, which helps to control the overall brain arousal and attention levels. However, in response to a stimulus or event, there is a rapid increase in ACh release, leading to phasic cholinergic activity (77). In this way,

we can induce phasic cholinergic activity with a warning cue, which results in a rapid release of ACh.

In the central nervous system, the ACh plays an important role in arousal, memory and other functions. However, the activation of the cholinergic system also produces peripherical alterations. In the periphery, the ACh can stimulate the parasympathetic nervous system since this system uses acetylcholine as its primary neurotransmitter (77,78). This system is responsible for promoting the “rest and digest” state.

1.4.2.3 Tasked-evoked pupil responses

In the state of alertness, the cholinergic and the noradrenergic systems, among others, interact, producing changes in both brain and body states to prepare for the upcoming stimulus. One example of this interaction is the pupillary response. Several studies have been developed to understand the mechanisms of task-evoked pupil responses (64,79). It is believed that the pupillary response results from the microstimulation of several structures, namely the LC, which triggers the activation of the noradrenergic system. Furthermore, it is postulated that the cholinergic system is also linked to the pupil's response, likely due to its close interconnection with the LC (80).

The activation of these two systems was also verified in (81). In an experiment with mice, the authors found high cholinergic and noradrenergic activation during pupil dilation. Additionally, they observed that, at the peak of dilation, there was a greater activation of noradrenergic activity. However, this noradrenergic activation rapidly decreased, whereas cholinergic activity persisted. That said, the authors found that that cholinergic activity has a more pronounced effect on the pupil diameter, allowing the pupil to remain dilated for longer. In this context, the noradrenergic system is presumed to assume a more pivotal role in the dilation process (correlating with pupil derivative), while the cholinergic system is responsible for sustaining the pupil dilated, that is, maintaining the pupil size.

1.4.3 Freezing state during anticipatory attention

As previously mentioned, when a warning cue is presented, the noradrenergic and cholinergic systems are induced. These two systems interact and work together to modulate alertness, ensuring that the individual is adequately prepared to the upcoming stimulus or threat (8). The activation of the noradrenergic and cholinergic systems stimulates both the sympathetic and parasympathetic branches of the autonomic nervous system. Nonetheless, this process is influenced by several factors, including the nature of the stimulus, its intensity,

the context, among other factors, that can affect how much each system is stimulated. For example, there are certain stimuli that produce a heightened response from the noradrenergic system, thereby producing a more pronounced activation of the sympathetic nervous system (8). In these cases, we have a state known as “fight and flight”, characterized by a significant increase in energy so that the body can react to the upcoming stimulus. On the other hand, there are certain stimuli, with different features, that cause an accentuated activation of the cholinergic system, which results in a higher activation of the parasympathetic nervous system. The predominance of the parasympathetic nervous system results in a state called “freezing state”, which is characterized by a behavioral inhibition (8).

During the freezing response, both the sympathetic and the parasympathetic nervous systems are activated. The physiological characteristics of freezing state encompass sympathetic features, induced by an increase in the neurotransmitter NE, and parasympathetic features, linked to the increase in the neurotransmitter ACh (8). The dominance of either system at a given moment determines the specific physiological profile observed. So, greater activation of the sympathetic system produces an increase in arousal and physiological changes, such as an increase in heart rate and blood pressure, inhibition of digestive function, enhanced respiration, among others. On the contrary, higher parasympathetic activation causes a deceleration in heart rate and inhibits fight-or-flight responses, which contributes to the overall state of freezing observed during threatening situations (8).

In summary, when a warning signal is presented, the state of expectation modulates cortical activity, stimulates the sympathetic system, which causes pupil dilation and changes in skin conductance, and activates the parasympathetic system, which results in a cardiac deceleration (82). Additionally, in this state, it is verified motor inhibition, which improves sensory processing and decision-making processes and facilitates threat assessment (7–9).

1.5 Classification algorithms

Machine learning has been increasingly used in neuroscience, enabling numerous advances in the analysis of neural data and, consequently, enhancing comprehension of how the brain works.

Machine learning uses mathematical algorithms to perform a given task. The algorithms receive a set of data that will be used as training to make estimates without being specifically programmed for this purpose (83–85). Thus, the main purpose of machine learning is to learn from data.

There are two main types of machine learning: supervised learning and unsupervised learning. In supervised learning, the algorithm uses labeled data in the training process (i.e., each example is associated to its label). Here, the goal of the algorithm is to learn a function that maps features vectors (input data) and corresponding labels (output categories). On the contrary, in unsupervised learning, the algorithm uses unlabeled data (i.e., only the data itself and not the corresponding labels), focusing on understanding the hidden patterns and structure of the data (83,85).

1.5.1 Support vector machine

A support vector machine (SVM) is a supervised learning approach for classification and regression analysis. As for others learning systems, the goal of SVM is to learn from a training data set and attempt to generalize and make correct predictions on unseen data (84–86).

For this, the algorithm identifies a hyperplane in an N-dimensional space (where N is the number of features) that effectively separates the data points into distinct classes based on examples in the training dataset. Thus, the optimal hyperplane is the one that has the maximum margin, i.e., is the one which is maximally distant from the labeled points of both classes located on either side. The maximum margin is the maximum distance between two points of different classes. The points lying closest to the decision surface are the ones with most influence on the position of the hyperplane and are called support vectors. The hyperplane is defined as $\mathbf{w}^{\mathbf{T}}.\mathbf{x}+b=0$, where b is the bias, $\mathbf{x}$ are the points located on the hyperplane and normal to it and $\mathbf{w}$ is a vector with the weights (84–86).

In this type of algorithm, it is possible to modulate the decision boundary so that it allows some classification errors to occur. In this case, we can define it as soft margin SVM. On the other hand, hard margin SVM looks for a hyperplane that accurately separates the data into two classes with no misclassification. To implement hard or soft margins, it is necessary to define a penalty parameter C, which controls the flexibility of the model, by defining the

classification errors that the classifier wants to avoid. In this way, this parameter allows the generalization of the data (84–86). For example, considering *C*=0 means that we do not want any misclassification and therefore we are facing the hard margin SVM. To maximize the margin, the optimization problem is the following:

$$minimize\ \frac{1}{2}\ \boldsymbol{w}^T\boldsymbol{w} + C\ \sum_{i=1}^{l} \xi_i$$

$$\boldsymbol{s.t.}\ \ y_i(\boldsymbol{w}^T\boldsymbol{x_i} + b) \geq 1 - \xi_i\ \ \forall i = 1, \dots, n, \xi_i \geq 0$$

Equation 1 *– Optimization problem* (86).

- C is the penalty parameter.
- ξ_i is the slack variable that measures, in each point, the violation of the margin restriction.

The SVM algorithm performs both linear and non-linear classification. The last type of classification is done using the kernel trick, a mathematical function used to map the inputs into high-dimensional resource spaces (85–87). That is, a kernel function projects data from a low dimensional space to a higher dimension space in order to make the problems linearly separable. To capture complex patterns and relationships in the data, the kernel trick computes the inner products between pairs of data points. There are several types of kernel functions commonly used, namely: linear kernel, polynomial kernel, radial basis function kernel (RBF) (85–87).

The RBF kernel is usually preferred over other kernels, because it can capture complex relationships between the input features and is more robust to overfitting. The RBF function determines the degree of similarity or proximity between two points X_1 and X_2 and can be defined by the following expression: (85–87)

$$K\ (X_1, X_2) = exp(-\ \frac{\|X_1 - X_2\|^2}{2\sigma^2})$$

Equation 2 *– RBF kernel.*

- σ is the kernel parameter and is the width of the kernel.
- $\|X_1 - X_2\|$ is the Euclidean distance between the two points.

The sigma parameter (σ) must be optimized, since this parameter controls the level of non-linearity introduced in the model. Beyond the sigma parameter, the penalty parameter *C* must also be optimized since this value introduces a penalty for misclassified data. There are several ways to optimize these parameters such as grid search (85–87), random search, and Bayesian optimization (88). All these optimization strategies attempt to minimize a loss (error) function computed on a separated set of data (named "validation" set) by varying the parameters. Matlab offers a dedicated functionality to perform such optimization, namely '*OptimizedHyperparamenter*' (89). The '*OptimizedHyperparamenter*' method uses Bayesian optimization, to find the optimal hyperparameters. Bayesian Optimization, based on Bayes' Theorem, is a technique used in global optimization problems, with the aim of achieving both efficiency and effectiveness (90,91). To do this, it is necessary to construct a probabilistic model to represent the objective function. This model, known as the surrogate function, is, subsequently, explored in order to select the candidate samples for evaluation on the actual objective function. The search for the optimum is done by the acquisition function, which is used to choose the point to be explored in the next iteration (90,91).

1.5.2 Deep learning

Deep learning is a branch of machine learning that exploits deep neural networks to solve the objective task (supervised or unsupervised), e.g., image recognition (92). Deep neural networks are composed by many layers of artificial neurons. Each of these layers learn simple non-linear functions during training. Thus, by composing these functions, in general a complex and non-linear function that maps the input to the output is learned. Depending on the connections established across neurons, different networks can be realized, e.g., recurrent neural networks, feed-forward and fully-connected neural networks, and feed-forward and convolutional neural networks (CNNs) (92).

1.5.2.1 Artificial neural networks

The structure of CNNs was inspired by the neurons in human and animal brains and is similar to a traditional neural network. To understand CNN, let us start by looking to the simplest neural network. This type of network has multiple layers of interconnected units, the neurons (83,93). We can divide the multiple layers into an input layer, one or more hidden layers and an output layer. Considering the input signals, a neuron can be activated, producing a new signal that is transmitted to another neuron. Each node, often referred to as an artificial neuron, is interconnected with other nodes and has a corresponding weight and threshold (83,93).

- **Weights** – the weights play a crucial role in determining the importance of a specific input. A higher weight denotes a greater influence of the corresponding input on the network's output, compared to other inputs. Each input's contribution is determined by multiplying its value by its respective weight. At the beginning, the network does not know the importance of the different inputs, so the weights are randomly assigned. Then, these weights go through an iterative adjustment process, allowing the network to capture the underlying patterns and relationships in the data, by assigning appropriate importance to different inputs (83,93).

- **Threshold** – A given node is activated if the output exceeds a defined threshold value. This node, when activated, sends the data to the next layer of the network. Conversely, if the output is below the mentioned threshold, no data is passed to the subsequent layer (83,93).

The activation of neurons in artificial neural networks can be summarized as follows: each neuron receives a weighted sum of its inputs. Following this, an activation function is applied to compute the output of that neuron. If the output exceeds a predefined threshold value, the neuron is activated, thereby passing information to the next neuron (83,93).

- **Activation function** – We can define activation function as a "rule" to determine how much the neuron will be activated. There are various types of activation functions, including: (83,93)
    - *Linear function* – the activation is proportional to the input. In fact, no activation function is applied to the sum of the inputs, which results in an output equal to this sum (Figure 1.6 – a)). (83,93)
    - *Sigmoid function* – receives any real value as input and produces an output in the range of 0 to 1. If the input is higher, the output tends to get closer to 1. On the contrary, a smaller input results in an output closer to zero. It is frequently applied to binary classification problems at the output layer (Figure 1.6 – b)) (83,93). A generalization of the sigmoid activation function for multi-class classification problems is the *softmax function*. The latter maps the input neurons' scores (e.g., N scores, corresponding to N classes) to a discrete probability distribution (over N classes).
    - *Hyperbolic tangent* – Similar to the previous function, the output is within a defined range, but in this case the range is -1 to 1. It is frequently employed in hidden layers and may record both high and low activation levels (Figure 1.6 - c)). (83,93)

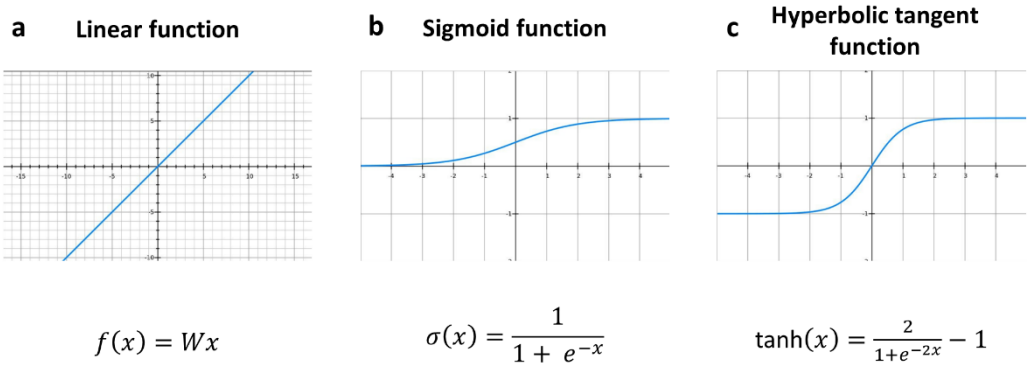


Figure 1.6 – *Activation functions. Adapted from* (93).

1.5.2.2 Convolutional neural networks

A Convolutional neural network is a class of deep learning networks, most commonly applied to tasks such as object recognition, image classification and text analysis. The architecture of the CNN is inspired by the organization of the visual cortex in animals, where groups of cells process different regions of the input image. To fill the complete visual field, these smaller subregions are tiled together. In this way, the CNN is mostly employed to extract features to a dataset similar to a grid-matrix (83,93).

The input data from the input layer is transformed by CNNs across all connected layers into a set of class scores provided by the output layer. Figure 1.7 shows a high-level view of a typical CNN organization.

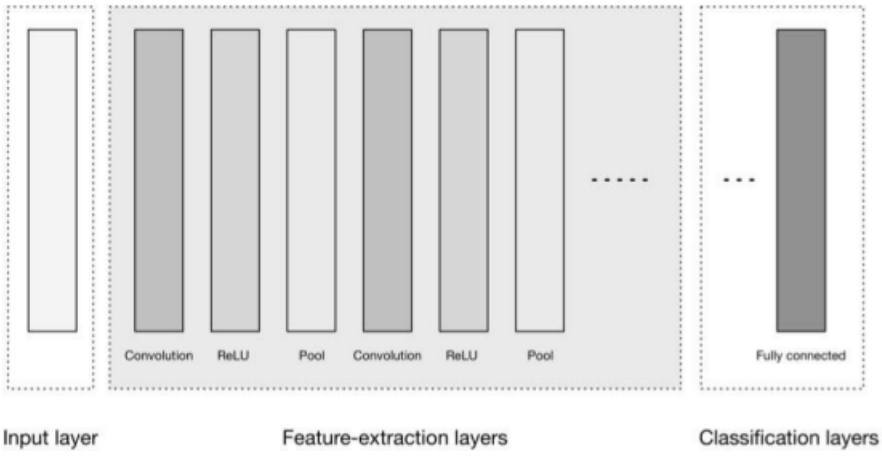


Figure 1.7 – *High-level CNN architecture. Adapted from* (93).

In Figure 1.7 we can observe three principal groups of layers. The input layer is where the input data is loaded and stored, which are then used by the second group to process the data. The features-extraction layers are essential for identifying and extracting relevant characteristics from the input data. Lastly, we have the classification layers that are responsible for taking the higher-order features extracted by preceding layers and producing class probabilities or scores. The classification layers typically include one or more fully connected layers (83,93).

Feature-extraction layers – this group of layers can include various types of layers, such as convolutional layers, ReLU layer, Pooling Layer, among others.

Convolutional layers – these layers are responsible for applying filters to the input data to extract meaningful features. These layers consist of a set of learnable filters, often referred to as kernels, that are convolved with the input data (83,93). The convolution process consists of multiplying a kernel with a specific region within the input image. In this way, each value is multiplied by the corresponding weight. The output of this layer is a matrix that stores the sum of all multiplications called feature map. A convolutional layer is composed of multiple filters, which are applied consecutively. This sequential application implies that once an input is fully processed by a particular filter, the network applies the subsequent filter. There are several parameters that characterize these layers, and thus, that alter the output of the convolution: (83,93)

- Filter size – represents the dimensions of the kernel array.
- Output depth – refers to the number of distinct filters or feature maps that a convolutional layer produces.
- Stride – determines how far the filter will move each time its function is applied.
- Zero-padding – allows us to obtain the desired dimension in the activation map, by adjusting the input size. This process consists of symmetrically adding zeros to the input matrix.

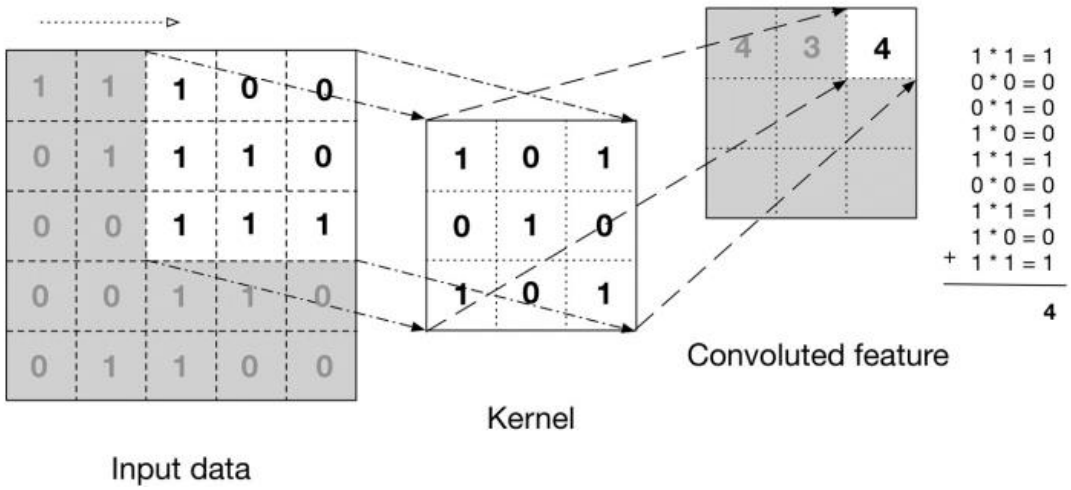


Figure 1.8 *– The convolution operation. Adapted from* (93).

Pooling layer – is typically added after convolutional layers, with the aim of downsampling the image and, consequently, reducing computation. The pooling layer achieves this by reducing the dimensions of the hidden layer by merging the outputs of neuron clusters from the previous layer into a single neuron within the subsequent layer. There are mainly two types of pooling operations: (83,93)

- Max pooling – in this type of pooling it is extracted the maximum value within each filter action region (Figure 1.9).
- Average pooling – this layer, in turn, applies the average value in the region caught in the filter.

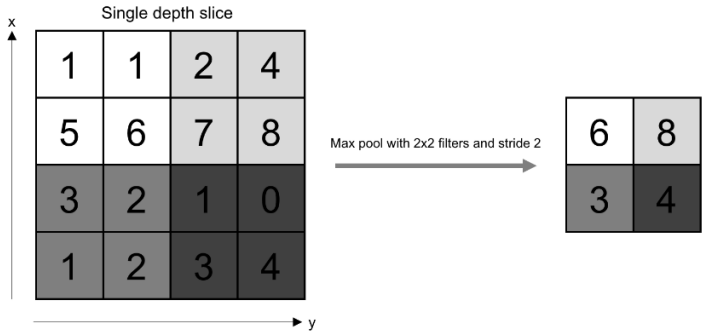


Figure 1.9 *– Max pooling layer example. Adapted from* (94).

ReLU Layer – in this layer it is applied an activation function, the ReLU function, defined by the expression $f(x) = \max(0, x)$, to the input x. If the input value is less than 0, this function returns 0, otherwise, it returns the input value. A ReLU layer introduces non-linearity to the algorithm, solving the problem related to the gradient vanishing during the training process (83,93).

Classification layers - The classification layers typically include one or more fully connected layers. This layer is used to compute the probabilities of the output classes for the input data. It takes the output of convolution/pooling operation, flattens the data and predicts the most appropriate label for the image. The output of this layer is created by multiplying the inputs with the associated weights, summing them, and then applying an activation function to produce the output. The results are propagated to the next fully connected layer, with the last layer containing a neuron for each class label. This last layer (properly activated via the softmax function) is responsible for producing the output probability distribution that represents the likelihood of each class label (83,93).

1.5.2.3 CNN's algorithm

Throughout the years, among the CNNs proposed for classifying EEG signals (95), EEGNet (96) represents the most used and successful CNN algorithm. This algorithm, developed by Lawhern V. et al. (96), is a compact CNN designed for classification and interpretation of brain-computer interfaces based on EEG signals. The architecture of EEGNet comprises two different blocks, each with its own purpose. The first block is responsible for learning how to filter in time and in space the input EEG. The second block, is responsible for learning how to resume in time the information (deep temporal feature learning), see Figure 1.10 (96).

In block one, two convolutional steps are sequentially performed. Initially, 2D convolutional filters are applied to generate feature maps that capture the EEG signal at different band-pass frequencies. Then, to learn a spatial filter, a *depthwise convolution* is used. The use of *depthwise convolution* is important since it reduces the trainable parameters. This operation allows learning a set of spatial filters for each temporal filter, facilitating the efficient extraction of the frequency-specific spatial filters. After that, the exponential linear unit nonlinearity is applied. To combat overfitting, given the limited sample sizes, dropout technique is employed. Finally, in block one, a pooling layer is applied to reduce the sampling rate of the signal (96).

In block two, the algorithm uses a *separable convolution,* to reduce the number of trainable parameters, and an average layer to reduce the dimension. Lastly, the features pass to a classification block, more precisely, a softmax layer. This layer is responsible for converting raw scores into probability distributions. These raw scores are normalized into a reliable probability distribution using the softmax method. Then, the predicted class is the one for which the softmax has produced the highest probability (96).

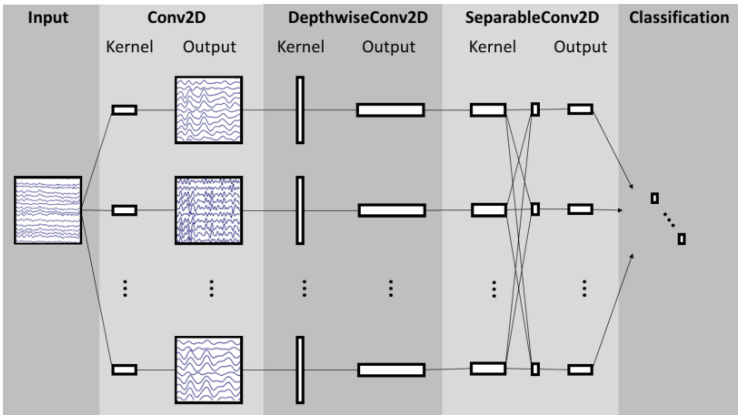


Figure 1.10 *- Overall visualization of the EEGNet architecture. Adapted from* (96).

1.5.2.4 CNN's algorithm optimization

Besides the parameters that characterize convolutional and fully-connected layers, the performance of CNN models also depends on several training options and methods. Here are described some typical training options:

- **Solver** – is responsible for model optimization, through generating parameter updates that aim to enhance the loss function. There are several solvers that can be used, for example, 'sgdm' that uses the stochastic gradient descent with momentum optimizer, or the Adam optimizer – Adaptive Moment Estimation optimizer (93,97).

- **Learning rate** – determine the step size at each operation during the optimization process in order to minimize the error of the algorithm's guesses (93,97)

- **Regularization** – is responsible for enhancing CNN's ability to generalize while preventing overfitting. There are several regularization methods: L1 and L2 regularization, dropout probability, among others. Coefficients L1 and L2 are used to prevent overfitting by making certain weights smaller. In turn, the dropout technique consists in randomly dropping out some neurons (93,97).

- **Mini-batch** – refers to a predetermined number of training examples. It is the number of examples used to estimate the error gradient before the update of the model weights (93,97).

- **Epochs** – represent the number of times that the training algorithm will iterate over the entire training set, i.e., one epoch is when the neural network processes the whole dataset only once (93,97).

- **Shuffle** – consists of data shuffling to help the flow of information. Here we also have a few options: 'once' – shuffle the data once before training; 'never' – do not shuffle the data; 'every-epoch' – consist in shuffling the data before each training epoch. (93,97)

- **Validation frequency** – Frequency at which a model is validated. (93,97)

1.5.3 Performance metrics

Performance metrics are used to measure effectiveness and quality of our classification model. Moreover, they provide insights about the capability of the model to predict what we want. To study the performance of the classification models there are a vast number of performance metrics that can be used, such as accuracy and area under Receiver Operating Characteristic (ROC) curve (AUC). The first measure (accuracy) tells us how many times the model has correctly predicted an item, i.e., consists in the division of the number of correct answers and the total of answers (85,98,99).

$$Accuracy\ (\%) = \frac{TP + TN}{TP + TN + FP + FN}$$

Equation 3 *– Accuracy expression.*

- *TP*: Examples that have been correctly classified as positive.
- *TN*: Examples that have been correctly classified as negative.
- *FP*: Examples that have been incorrectly classified as positive and are therefore actually negative.
- *FN*: Examples that have been incorrectly classified as negative and are therefore actually positive.

Models often use data that is not balanced, i.e., the number of instances in different classes is significantly different. In imbalanced datasets, a model can achieve high accuracy by simply predicting the majority class for all instances, ignoring the minority class. For

instance, if 95% of instances belong to class A and 5% belong to class B, a model that predicts all instances as class A will achieve 95% accuracy. However, this does not reflect how well the model can predict the minority class. This problem is addressed by balanced accuracy by considering the performance in each class (100).

Balanced accuracy is based on two metrics: sensitivity and specificity. Sensitivity is the ability of the model to correctly identify positive classes. On the contrary, the specificity is the model's ability to correctly identify negative cases (100).

$$Sensitivity = \frac{TP}{TP + FN}$$

Equation 4 *– Sensitivity expression.*

$$Specificity = \frac{TN}{TN + FP}$$

Equation 5 *– Specificity expression.*

$$Balanced\ accuracy\ (\%) = \frac{Sensitivity + Specificity}{2}$$

Equation 6 *– Balanced accuracy expression.*

The AUC is a widely used performance metric for binary classification models. It evaluates the model's ability to distinguish between positive and negative instances at various classification thresholds. The ROC curve is a graphical representation of the true positive rate (sensitivity) against the false positive rate (1 – specificity). Each data point on the ROC curve corresponds to a unique pair of sensitivity and specificity values, which is associated with a particular decision threshold (85,98,99). By calculating the area under the ROC curve, the AUC summarizes the model's overall predictive performance. A higher AUC value indicates a better-performing model that can effectively differentiate between positive and negative instances across different thresholds (85,98,99).

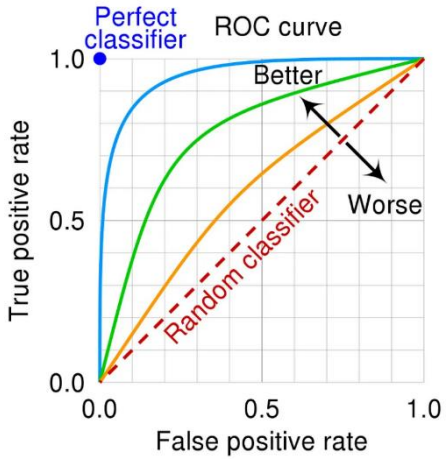


Figure 1.11 *– The ROC curve for different classifiers. Adapted from* (99).

1.6 Current study

In this study, we first analyzed data from a previously acquired dataset. In parallel, we adapted the previously used behavioral task and acquired data from a new set of participants using the improved version of the task. In both tasks, visual stimuli from two categories (cars and houses) were presented. Before the stimuli, a warning auditory cue was presented to signal the start of the trial and evoke a **state of attentive anticipation**. The first experiment used a visual task where the visual stimuli were made difficult to perceive via a backward masking procedure - backward masking study. In the second experiment, the stimuli visibility was modulated by using low image contrast - low contrast study. Moreover, in the second study, the participants reported not only the stimulus category they perceived but also if they recognized the stimulus or if they were responding by chance. The task and stimuli details are depicted in the Method's section.

In the backward masking study, the stimulus was followed by a mask. However, the question arose regarding the participant's actual perception of the stimulus. It was observed that when the participants did perceive the stimulus, many of them reported seeing only the cars, so they often only selected the houses if they did not see the cars. That said, during the analysis of the data obtained from this task, the doubt arose whether the difference in behavior was effectively between the difference in stimuli (car or house) or whether it was the difference between seeing a car or not seeing it. Another question that arose was whether, many times, they did not see the stimulus, but only the mask, reporting only the "object" they saw in the mask. Allied to this, we can also consider as a hypothesis that fluctuations in

cortical sensitivity may lead to improved perception of the mask but hinder the perception of the stimulus, potentially making it difficult to interpret the results. Based on this, we decided to change the task. So, we decided to remove the mask and lower the contrast of the presented stimulus making it difficult to recognize. This type of approach was based on the approach developed by Podvalny, E. et al. (10). Our expectation in changing the approach was to design a task hard to accomplish in which the participant would only recognize 50% of the trials.

In addition to changing the stimulus presented, we also implemented, in the second version of the task, a new feature where the participant had to report whether or not he had recognized the stimulus, allowing us to determine the participants' recognition of the stimulus. When given the two options (car or house), the participant, even without having seen the stimulus, has a 50% probability of getting its category right, that is, many of the trials defined as correct result from a random choice by the participants. Given this, in the backward masking study, numerous trials were erroneously attributed to the category of correct trials, which can bias the results and, consequently, contribute to the noise within the study. In this way, we believed that this implemented feature significantly enhanced our comprehension of stimulus detection. Consequently, we believe that the study of stimulus detection, instead of trial performance, affords a more realistic way to explore the impact of pre-stimulus activity, since we know exactly whether the participant perceived or not the stimulus. This way, and given we have this additional information about stimulus detection, we decided, in the low contrast study, to focus our analyses on the recognition of the stimulus, rather than the performance of the trial (whether or not the participant got the category right) as in the backward masking study.

The methods for these studies are described next.

2 Methods

In this chapter, we will start by describing the behavioral task that was used, as well all of the changes that were made to it throughout the project. The methods used to acquire and analyze the data were then explained.

2.1 Participants

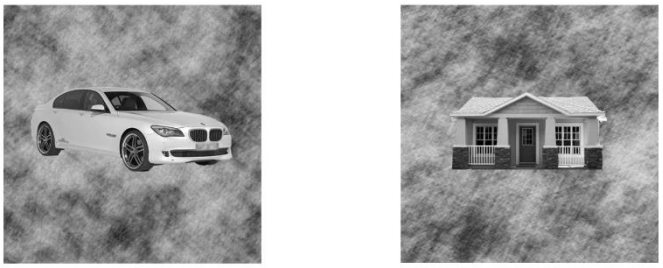
Forty-one participants took part in the present study: 18 in the backward masking study and 23 in the low contrast study. Some participants ended up being discarded due to problems in data acquisition or due to changes in task parameters. So, after all, we had 35 eligible participants, 16 from the backward masking study and 19 from the low contrast study. The information about our population is located in Table 1.

	Age (mean ± SD)	**Sex**	**Education** (mean ± SD)	**Dominant hand**	**Dominant eye**
Backward masking study	24.72 years ± 3.71 years	12 women 6 men	17.22 years ± 2.95 years	right: 15 left: 3	right: 10 left:8
Low contrast study	25.74 years ± 6.39 years	21 women 2 men	16.74 years ± 2.61 years	right: 21 left:2	right: 11 left:12

Table 1 - *Participant's characterization. SD: standard deviation*

2.2 Stimulus

In this project, we used a discrimination visual task developed in (101), where the participants were asked to discriminate visual stimuli between two categories: cars and houses (Figure 2.1). The stimuli used were adapted from the fLoc functional localizer package from (102). The authors selected these stimuli since they activate distinct regions of the visual cortex. In particular, the presentation of a house results in a higher activation of the parahippocampal place area since this area is associated with the visual processing of buildings and places. In turn, car stimuli lead to higher activation of the fusiform gyrus area (101,103–105).



a b

Figure 2.1 *– Example of the two types of stimuli (**a** – car stimulus: **b** – house stimulus).*

In the backward masking study, the visual stimulus was difficult to perceive due to the use of a mask (Figure 2.2 – a)), which was composed of parts of houses and cars, adapted from the fLoc functional localizer package (102).

In the low contrast study, instead, we used visual stimuli at the threshold of subjective recognition through the manipulation of their contrast levels. To change the contrast of the visual stimuli, we based on the approach used in (10). The pixel intensities, ranging from 0 (black) to 255 (white), were normalized for the range of [0,1] and the mean was removed to centralize the values around zero. Then, the chosen contrast was applied to the image by multiplying it with a coefficient *c* (represents the contrast level). Finally, the background color of the image was then applied. Summarizing, the image was normalized and subjected to a number of operations to transform it based on the contrast and background coefficients. It is important to note that this approach was employed on both visual stimuli and scrambled images.

Allied to this, in the second experiment, we incorporated trials with scrambled images (without any object stimulus) (Figure 2.2 – b)), taken from the fLoc functional localizer package (102), to investigate whether participants had understood the task mechanisms linked to the recognition of visual stimuli.

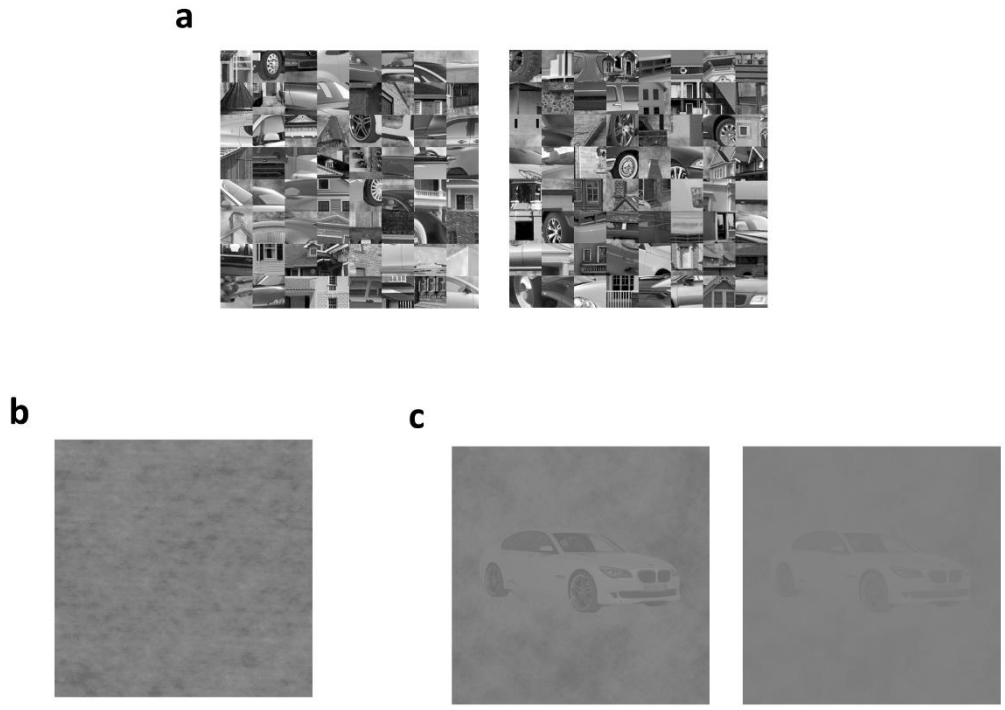


***Figure 2.2** – **a** Example of two types of masks. **b** Image used as scrambled image. **c** Example of two visual stimuli with different contrasts.*

2.3 Behavioral task

The task was developed in (101) using the version 2021a of Matlab and was updated with the version 2022a. To develop the task, it was necessary to use the *Psychtoolbox-3* (101,106). This toolbox is composed of a set of functions for vision and neuroscience and research, facilitating the synthesis and presentation of visual and auditory stimuli as well as interaction with the viewer (101,106).

For each participant, we acquired around 40 minutes of task, divided in 10 min runs. Each run consisted of 60 trials, in the backward masking study, and 65 trials, in the low contrast study. Four participants from the backward masking study and three from the low contrast study performed an extra task run with a total of five runs.

Each trial started with an auditory warning cue that alerted participants of the upcoming stimulus. The interval between the cue and the subsequent visual stimulus was randomized between two and six seconds, to prevent participants from accurately predicting the timing of the visual stimulus display, inducing a state of expectation. Then, the visual stimulus was presented for 30ms. After the presentation of the visual stimulus and the mask, in the backward study, a response prompt was presented to the participants. To ensure that

the motor response did not impact the EEG data during stimulus processing, the response sides associated with the options of the response prompt were mapped to left/right or right/left randomly for each trial.

- **Backward masking study** (Figure 2.3 – a)

In this study, the visual stimulus was followed by a visual mask. Then, after the presentation of the response prompt, the participants were instructed to indicate which stimulus was presented by pressing a button using either their left or right index fingers (keys 'Z' or 'M'). That said, in this task, we were only able to extract accuracy for a given trial rather than stimulus recognition.

- **Low contrast study** (Figure 2.3 – b)

In our second study, we changed the contrast of the visual stimulus so that it was presented at the threshold of subjective recognition. To ensure that the image contrast was adjusted to reach a 50% subjective recognition rate for each individual participant, we implemented an adaptive threshold procedure.

The participants, in this version, were instructed to report whether they recognized the stimulus, allowing us to be aware of stimulus recognition in each individual trial. If they had actually seen any object, i.e., if they recognized the stimulus, they responded with the right and left middle fingers and pressed the keys 'Z' or 'M' based on the stimulus that they saw. If they did not see any object, i.e., if they did not recognize the stimulus, they must guess which stimulus was being presented and, to measure that, they were instructed to answer with the right and left index fingers. So, when they did not see the stimulus, they had to answer with the 'X' key or with the 'N' key.

Lastly, in this low contrast study, scrambled images were added, in order to verify that the participant was correctly following the instructions in cases where he/she recognized and in cases where he/she did not recognize the stimulus. That is, we expected that, in the trials in which the scrambled images were presented, participants would be unable to recognize any specific object and consequently would consistently respond using the left and right index fingers.

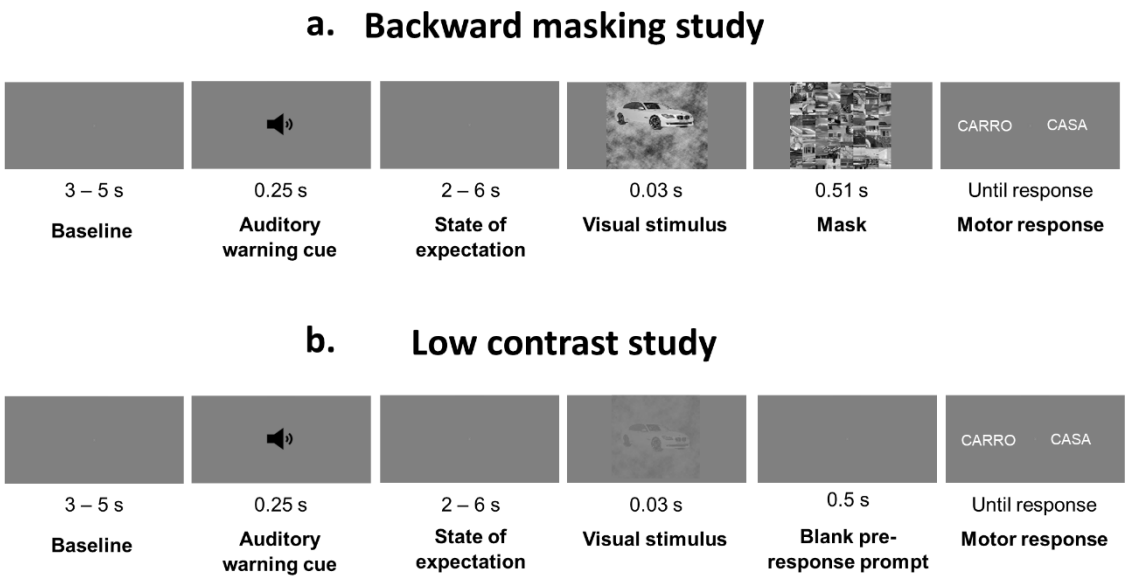


Figure 2.3 *- Schematic example of one trial type (**a** - backward masking study; **b** - low contrast study) (the response prompt, and the visual stimulus here are not in scale – the size was increased to facilitate visualization).*

To aid data analyses, trigger pulses were generated at the onset of each stimulus and at every button press. Considering the approaches adopted for each study, certain events were included in the low contrast study that were not present in the backward masking study.

Backward masking study

EEG and Biopac event code	Eye tracker message	Event
1	'SUBJECT'	Start of acquisition
2	'sound'	Auditory cue
3	'estimuloCASA'	"House" stimulus (30ms of duration)
4	'estimuloCASA'	"House" stimulus (40ms of duration – only for subjects 3 and 4)
5	'estimuloCASA'	"House" stimulus (50ms of duration – only for subjects 3 and 4)
6	'estimuloCARRO'	"Car" stimulus (30ms of duration)
7	'estimuloCARRO'	"Car" stimulus (40ms of duration – only for subjects 3 and 4)
8	'estimuloCARRO'	"Car" stimulus (50ms of duration – only for subjects 3 and 4)
50	'mask'	Mask
9	'OptionLEFT_CASA'	Response prompt with "house" option on the left
10	'OptionLEFT_CARRRO'	Response prompt with "car" option on the left
11	'Response_z_hit'	Response: participant chooses the option on the left correctly
12	'Response_z_miss'	Response: participant chooses the option on the left incorrectly
16	'Response_m_hit'	Response: participant chooses the option on the right correctly
13	'Response_m_miss'	Response: participant chooses the option on the right incorrectly
14	-	End of experiment

Table 2 *– Triggers and the corresponding events of the backward masking study.*

Low contrast study

EEG and Biopac event	Eye tracker message	Event
1	'StartAquisition'	Start of acquisition
2	'sound'	Auditory cue
3	'estimuloCASA'	"House" stimulus
6	'estimuloCARRO'	"Car" stimulus
60	'estimuloSCRAMBLED'	"Scrambled" stimulus
9	'optionLEFT_CASA'	Response prompt with "house" option on the left
10	'option_LEFT_CARRO'	Response prompt with "car" option on the left
11	'Response_z_correct'	Participant chooses the option on the left correctly – recognized the stimulus
12	'Response_z_wrong'	Participant chooses the option on the left incorrectly – recognized the stimulus
16	'Response_m_correct'	Participant chooses the option on the right correctly – recognized the stimulus
13	'Response_m_wrong'	Participant chooses the option on the right incorrectly – recognized the stimulus
20	'Response_x_correct'	Participant chooses the option on the left correctly – did not recognize the stimulus
21	'Response_x_wrong'	Participant chooses the option on the left incorrectly – did not recognize the stimulus
22	'Response_n_correct'	Participant chooses the option on the right correctly – did not recognize the stimulus
23	'Response_n_wrong'	Participant chooses the option on the right incorrectly – did not recognize the stimulus

61	'Response_scrambledz'	Participant chooses the option on the left in scrambled trials – recognized the stimulus
62	'Response_scrambledm'	Participant chooses the option on the right in scrambled trials - recognized the stimulus
63	'Response_scrambledx'	Participant chooses the option on the left in scrambled trials - did not recognize the stimulus
64	'Response_scrambledn'	Participant chooses the option on the right in scrambled trials - did not recognize the stimulus
14	'EndAquisition'	End of experiment

Table 3 *– Triggers and the corresponding events of the low contrast study.*

2.4 Adaptive threshold procedure

In the low contrast study, we implemented an adaptive threshold procedure (Figure 2.4), wherein the image contrast was adjusted to reach a 50% subjective recognition rate for each individual participant. The procedure consisted of successively presenting a predetermined set of stimuli with different contrasts. This set of stimuli was repeatedly displayed in a randomized order, ensuring equal frequency of occurrence for each contrast level. In each trial, the participant reported whether he recognized or not the stimulus by pressing different keys. For this, they had to respond with left or right middle fingers if they recognized the stimulus and with left or right index fingers if they were just guessing the category, that is, if they had not recognized the stimulus. The proportion of responses was calculated for each specific contrast value, after the presentation of all contrast levels multiple times. Finally, to determine the contrast level corresponding to the 50% subjective recognition rate, we used a psychometric function.

During the first acquisitions, we noticed that the contrast that represents 50% of subjective recognition rate was not reflected in 50% of recognition of the trials during the task. This fact can have several causes. First, the participants may not have understood the task instructions and even having recognized the stimulus, they might have answered with their index fingers (responses associated with non-recognition of the stimulus), because, for example, they were not completely sure of the stimulus category. Second, maybe the number of stimuli used in each contrast was not enough to determine, in an accurate way, the contrast that represents a 50% subjective recognition rate. Based on that, we decided to double the number of stimuli in each contrast from participant 26. With this improvement, it was possible to reach a recognition closer to 50% when performing the task.

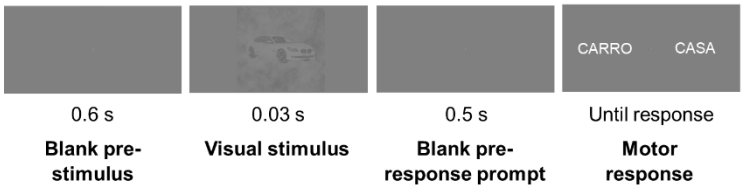


Figure 2.4 *– Schematic example of one trial type from the adaptive threshold procedure (the response prompt and the visual stimulus here are not in scale – the size was increased to facilitate visualization).*

2.4.1 Psychometric Function fitting

Psychometric functions allow us to relate a certain participant's psychophysical performance, as the recognition rate of a stimulus, and specific physical stimulus characteristic. This type of measure is done to determine parameters that capture the participant's behavior. To obtain the parameters, a continuous function is fitting to the data. The process to fit the curve is done iteratively through a range of possible values of two parameters, alpha and beta (107,108).

We used the following formulation of the psychometric function:

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

Equation 7 *– Psychometric function*

The x parameter refers to stimulus intensity and the ψ refers to a measure of performance (in this thesis, the measure that we used was the proportion of recognized stimuli). The α value determines the curve's general position along the abscissa relating to sensory or perceptual processes. The parameter β determines the slope or gradient of the curve. Lastly, we have the γ and λ parameters. the parameter λ is the lapse rate and corresponds to the proportion of trials where observers will respond independently of stimulus level, for example, represents the probability of responding incorrectly because of a lapse. The parameter γ, in turn, represents the guessing rate, which denotes the chance-level performance (108,109).

To model psychometric data, we choose the logistic function. This function is given by:

$$F(x;\ \alpha,\beta) = \frac{1}{1 + \exp\,(-\beta(x - \alpha))}$$

Equation 8 *– Logistic function expression.*

After choosing the psychometric function and the initial parameters values, we can move forward with, iteratively, fitting the logistic curve to the data. The fitting involves searching for values that produce a curve that closely aligns with the experimental data (107,108). To find the best fitting of the function to our data, we used Maximum Likelihood criterion. Using this criterion, we were able to determine which psychometric function best recreated the experiment as if it was carried out by a subject (110).

Considering the characteristics of the mentioned parameters, we have determined that γ and λ parameters are considered "fixed parameters", since they do not change during the entire process. In opposite to this, α and β vary during the fitting, so they are classified as "free parameters" (110).

The purpose of the use of these functions is to find the contrast that allows a recognition of 50% of the stimuli. Consequently, the α parameter will fluctuate between the contrast values that have the recognition rate close to 0.5 (we used the fraction and not the percentage). Then, we have the β parameter, which is difficult to estimate by inspection. Due to the results, we expect that there is a large variation in the recognition rate over a small change in contrast levels, so we put a high number here, like 50 (110). The γ and λ, in turn, were considered zero.

To evaluate the errors in the estimated parameters, we used the Standard Error. This method determines how far the estimated values will probably deviate from the true value. The standard errors are estimated using a method called bootstrap analysis, which generates 400 hypothetical sets of data. The logistic function is used to fit each hypothetical data set to estimate α and β. The parameter errors are subsequently determined by calculating the standard deviations of these predicted values across all sets (107,109,110).

Finally, to measure how well the model fits the data, we used the goodness-of-fit of the function. To do that, we employed the method described in (111). The method involved generating a Monte Carlo distribution from the real data, creating multiple artificial datasets through randomization. After that, the method determined the percentage of simulations with higher deviation than the original data, which allows us to extract an associated p-value. The p-value ranges from 0 to 1, with a higher the value indicating a better match between the fitted function and the data (107,109,110).

One of the toolboxes that allow us to implement these functions is the *Palamedes* toolbox. This toolbox is characterized by a collection of MATLAB functions and demonstration programs for the analysis of psychophysical data (108).

2.5 Physiological recordings: EEG, ECG, respiration and eye tracking

As was said before, the aim of our study was to analyze how pre-stimulus activity modulates cortical processing of visual stimuli and, consequently, visual perception. So, while participants were engaged in the task, we simultaneously acquired EEG and body physiological signals.

The EEG signal was acquired with a 64-channel Neuroscan system with scalp electrodes placed according to the International 10–20 electrode placement standard, with reference between the electrodes CPz and Cz and ground between FPz and Fz. Acquisition rate was 500 Hz. Vertical and horizontal electrooculograms were recorded to monitor eye movements and blinks. This system also allowed us to measure other signals using bipolar electrodes:

- Cardiac activity by placing the positive electrode in the right shoulder and the negative electrode over the lower end of the sternum.
- Blinking through electrooculography with electrodes placed on the face.
- Trapeze muscle activity using bipolar electrodes placed in the right shoulder (these data were acquired but not analyzed within the scope of this work).

To study respiration, we measured the circumferential changes in the thoracic region of the torso with a transducer belt with the Biopac system. Until subject 22, breathing data was recorded using a sampling rate of 5000Hz. Subsequently, the sampling rate was lowered to 1000Hz. We also measured leg movements with accelerometers with Biopic's System Bionomadix and, for the low contrast study only, the cardiac sound with the Biopac physiological microphone (these data were acquired but not analyzed within the scope of this work).

Finally, to study eye movements and pupil dilation, we used the EyeLink 1000 Plus. Acquisition rate was 500 Hz. During the data acquisition process, there were changes in the methodology used to study eye movements and pupil dilation. Initially, we used a monocular method, that is, the ET recorded the movements of only one eye, the dominant eye. From subject 30, we started using a binocular method that captured the movements of both eyes. The monocular method was used in 29 participants and the binocular method was used in the remaining participants, 12 participants. In the binocular method, to study the pupil, we selected the behavior of the pupil referring to the dominant eye. Blinks and saccades were identified as such when detected in both eyes.

To summarize, in our project, we used the EEG and ECG recordings, respiration, pupil dilation, blinking and saccadic activity.

2.6 Data correction

In the participants between S03 and S07 from the backward masking study and S21, S23 and S25 from the low contrast study, the events from the biopac system (respiration) and Neuroscan (EEG and ECG) were not correctly sent, with some events missing. Considering that the events from the ET were well acquired it was only necessary to import the event data from the ET into Matlab.

The ET and the other signals started recording at different times. So, it was necessary to measure the difference between the first and the second event from each trial, for both the ET and the remaining signals. When we discovered a trial where the difference is the same for the ET and for the other signals, means that in that trial the events are correct. Based on that, we determine the difference in latencies from the signals (specifically in that trial) and apply that difference to all events from the ET. Basically, we add the events from the ET to the respiration, EEG and ECG signals, but with the correct latencies.

2.7 Data analysis

To analyze all the data, we used Matlab (version 2022b) custom scripts and EEGLAB toolbox (version 2022.1) (112). In this section, for each physiological recording, we will describe the signal processing and subsequent analysis methodologies.

As mentioned, one of the goals of our project was to investigate how the state of expectation produced by the warning cue was reflected in changes in body and neural physiological signals and if the pre-stimulus brain and body activity modulated visual perception. So, in our analyses, we started by studying each signal individually. In each physiological signal we follow the following steps:

- Evaluate how body signals vary throughout the state of expectation.
- Evaluate the differences in body and brain signals between correct and incorrect trials – backward masking study - and between recognized and not recognized trials – low contrast study.

Reaction time: In addition to comparing the trial types analyzed in the two versions of the task, we also correlated the different features of each physiological signal with the RT. We decided to study the RT, as we believe that this measure more accurately reflects the influence of the state of expectation than the trial type in the backward masking study. This belief relates to the fact that many trials were incorrectly attributed to the group of correct

trials, as described above. It should be noted that comparisons involving reaction time were exclusively conducted using body activity. Given that EEG is composed of many channels, the study of each channel individually would be difficult, so we decided to discard these analyses and focus only on the relationship between reaction time and body physiology.

- To assess the reaction time, we considered the temporal distance between the presentation of the response prompt and the moment in which the participant pressed the key.
- The comparison between the reaction time and each body signal was carried out on a trial-by-trial basis through the correlation of both parameters. This approach was performed for all features of the various physiological signals, except for the phase of the cardiac and respiratory cycles in which the stimulus is presented. For these two features we used a regression model (described in the statistical analysis section).

Classifiers: After that, we integrated the pre-stimulus neural and physiological data in the classification models to predict participant's visual performance. Our expectation was to identify patterns in the different physiological measures that could facilitate the prediction of stimulus detection. Lastly, the final analyses aimed to determine whether the pre-stimulus body and neural activity modulated the way in which the stimulus was decoded. This classifier was trained using single-trial EEG data, with the objective of discriminating the category of the trial.

2.7.1 Neural activity

2.7.1.1 Preprocessing of EEG data

The EEG preprocessing was made using the *EEGLAB toolbox* (112). First, the EEG signal was re-referenced to linked earlobes and then bandpass filtered between 0.5-100Hz. After that, bad channels were removed and replaced through interpolation. Through visual inspection, some intervals of the data in which a generalized noise was seen, caused, for example, by the fact that the participant had moved his head, were also removed.

Then, the ICA algorithm was applied and the components that did not have a neural origin were removed. This process was conducted through visual inspection considering diverse features of the various components, such as the topography of the component, signal frequencies, the amplitudes, among others (112,113). Through the ICA algorithm, we can isolate various categories of signals (113). For example, when examining components that exhibit regular dipoles with smooth topography and that peak at physiological frequencies, it

is likely that these components originate from neural sources (113). In turn, if the components have stability intervals with occasional and very quick transitions, probability this component has origin on the saccades (113). Components originating from muscular activity are also easy to detect, since they are generally very focal and cover a local group of electrodes. Allied to this, in the data representation, we can see constant noise that does not vary with task events (113). Finally, this algorithm also allows us to eliminate components originating from bad channels. This type of component shows strong amplitudes, which are not correlated with other channels (113).

A significant number of the intervals that were removed contained the event corresponding to the stimulus presentation. Thus, when these intervals were removed, the associated event was also eliminated, making it impossible to apply epochs to these particular trials. In the following table, we show the number of trials used in the classifier for each subject.

Backward masking study

Subject	Runs	Total trials	Trials used
3	5	300	180
4	5	300	299
5	4	240	240
6	4	240	240
7	4	240	239
8	4	240	231
9	4	240	239
10	4	240	239
11	4	240	205
12	4	240	231
13	4	240	230
14	4	240	235
15	4	240	240
16	5	300	300
17	4	240	237
18	5	300	284

Table 4 *- Overview of total trials and trials that were included in the analyses in the backward masking study.*

Low contrast study

Subject	Runs	Total trials	Total trials with stimulus	Trials used
21	4	260	240	236
22	5	325	300	290
23	4	260	240	238
24	4	260	240	234
25	5	325	300	299
26	4	260	240	239
27	4	260	240	240
28	5	325	300	299
29	4	260	240	240
30	4	260	240	240
31	4	260	240	240
32	4	260	240	240
33	4	260	240	239
34	4	260	240	237
35	4	260	240	239
37	4	260	240	240
38	4	260	240	239
39	4	260	240	239
40	4	260	240	234

Table 5 *- Overview of total trials and trials that were included in the analyses in the low contrast study.*

2.7.1.2 Analysis of EEG data

First, we studied the effect of the cue on the EEG signal by studying the ERP evoked by the cue in an epoch between 0.2s before the cue and 3s after the cue onset (with baseline removed) and investigated where the ERP was significantly different from zero.

Then, we also compared the ERP across trial types. Since we wanted to evaluate the differences in neural activity just before stimulus onset, we ended up studying the signal just two seconds before and two seconds after the stimulus presentation. We chose to include two seconds after stimulus onset, in order to analyze the ERPs evoked by the target across the trial types considered. In this type of analysis, we had only the representation of each electrode individually; however, we wanted a way to represent all channels. For that, we used the *topoplot* function from EEGLAB in Matlab. This function allows us to plot a topographic map of an EEG field through interpolation. Since *topoplot* receives as input a value to each channel, we started by obtaining the mean value of each channel for each participant and then applying the average value of all participants. Since we only wanted to compare pre-stimulus activity across trial types, we extracted the average ERP amplitude in the one second just before target onset. For visualization of scalp topographies, we employed perceptually uniform color maps by Crameri et al. (114).

We also studied the EEG signal evoked by the stimulus, the ERP. To do that, we epoched the data in a time window of 200ms before and 500ms after the stimulus onset and removed

the baseline. We chose 500 ms since the response prompt only appears 500 ms after the stimulus and therefore this time window is free from motor processing signals.

Finally, EEG activity was included in the classifier. In the first analysis, in which the objective was to use pre-stimulus neural activity to predict visual performance, we used the continuous EEG signal. Next, we also used the continuous signal to discriminate the trial's category. In this case, we used the EEG signal after the presentation of the stimulus. Considering that one of the goals was to combine the produced ERPs with pre-stimulus activity, we decided to summarize the EEG pre-stimulus activity through a parameter for each channel of the EEG signal. To do this, we extracted the average amplitude in the one second just before target onset.

2.7.2 Cardiac activity

2.7.2.1 Preprocessing of cardiac data

The dataset of all participants was processed as described in (115). To process the data, the *EEGLAB toolbox* and the *ECG toolbox* from the Health Informatics Lab at Centro de Informática e Sistemas da Universidade de Coimbra (116) were used.

The ECG toolbox was specifically designed to handle data with a sampling rate of 250Hz. Consequently, we had to downsample the data from 500Hz to 250Hz.

To investigate the influence of the cardiac cycle in the state of expectation, we decided to measure the heart rate at the moment of stimulus presentation, as well as monitor heart rate fluctuations during this particular state. In order to compute the participant's heart rate accurately, it was necessary to detect the peaks, using the *ECG toolbox*. However, as the toolbox was originally designed to work with ECG data obtained using the precordial lead II, it unfortunately misidentifies many of the peaks. To address these limitations, we used Matlab scripts developed by (115).

In most participants, the S peak was very prominent and the one that was correctly identified more often (Figure 2.5). Therefore, we used the S peak latency to calculate heart rate throughout the task.

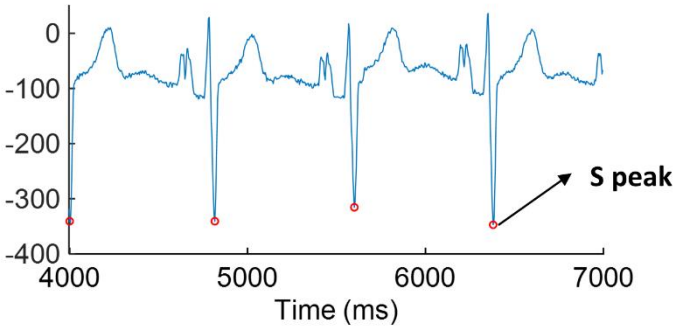


Figure 2.5 *– Example of a typical ECG data, where the red circles denote the S peaks.*

In this way, we used the S peaks to measure the heartbeat. Initially, we determined the time interval between successive S peaks within each cardiac cycle. Then, having the length of each cardiac cycle, we converted them, which were originally expressed in milliseconds, into heartbeats (beats per minute). This conversion was achieved by dividing 60000 by the cycle duration of each cardiac cycle.

Due to problems in acquiring the ECG signal or even in detecting peaks, we sometimes obtained ectopic beats, that is, with abnormal durations. Therefore, to exclude these beats, we calculated the z-score of the duration of each beat and removed values with absolute z-scores greater than five. It should be noted that, when obtaining the instantaneous heart rate, we excluded specific values and not trials associated with ectopic beats. This process was done before *spline* interpolation. However, in the analysis in which the duration of the cardiac cycle was considered, if the cycle before the stimulus or before the cue (these cycles are used in this discrete analysis) were considered artifacts, the trial would be removed and would not be subject to analysis.

After that, the values of heartbeat extracted in the previous step were interpolated using function *spline* in Matlab. This function enabled us to compute, at each time point, the instantaneous heart rate.

2.7.2.2 Analysis of cardiac data

First, we studied the effect of the cue on cardiac activity, more precisely in the heart rate. For that, we used the interpolated heart rate that assigns a value of heart rate for all data points. To study the variations in heart rate induced by the cue, we used an epoch between 2s before the cue and 10s after the cue onset (with baseline removed) and investigated the time points where the heart rate variation was significantly different from zero. To evaluate the heart rate variation just before stimulus onset and to study how its presentation influenced

cardiac activity, we ended up studying the signal just two seconds before and three seconds after the stimulus presentation. We also compared the heart rate variation across trial types, using the same time windows, and evaluated the time points where the difference was statistically significant.

Then, we analyzed cardiac activity by comparing cardiac parameters between correct/incorrect trials or recognized/unrecognized trials and studied the correlation between cardiac parameters and reaction time. The extracted cardiac parameters were then used as features in the classifiers.

For these analyses, we decided not to use the interpolated heart rate, since it is an approximate measure, and focus only on discrete heart rate. The discrete heart rate technique allowed us to extract heart rate values at each heartbeat. For the discrete heart rate values, we used the duration of the cycle immediately before stimulus presentation and the relative heart rate, which represents the heart rate variation between the cardiac cycle immediately before the auditory cue and the duration of the cardiac cycle immediately before visual stimulus onset. We decided to select the cardiac cycle immediately preceding the stimulus presentation, as we believe that the cycle that contained the stimulus would already be influenced by the processes associated with the processing of the stimulus.

Moreover, we also studied the effect of cardiac phase in visual perception. To do that, we estimated the phase of the cardiac cycle at the moment of visual stimulus onset by dividing each cycle in 360º and calculating the phase (in degrees) at that moment. As the phase is a circular variable, to study the impact of cardiac phase, we studied the impact of the sine and cosine of the phase together and separately. This approach takes into account the circular component of the cycle by projecting it into cartesian coordinates.

2.7.3 Respiratory activity

2.7.3.1 Preprocessing of respiratory data

The dataset of the participants was processed as described in (101). In this previous study, the respiration data from the backward masking dataset was analyzed. We based our preprocessing of these data on their analyses. Initially, the data was downsampled and filtered using the *EEGLAB toolbox*. Considering that the typical breathing rate, in healthy adults, ranges from 12 to 20 breaths per minute (117), the data was downsampled to 100 Hz. Then, the data were filtered with a bandpass filter limited between 0.01-2Hz. Even after filtering, the data continued to have a significant amount of high frequency noise. So, we smoothed the data with the '*loess*' method.

Analyzing a typical breathing data (Figure 2.6), we can see that it consists of peaks and valleys. A peak corresponds to the end of an inhalation and the beginning of an exhalation. A valley, in turn, corresponds to the end of an exhalation and the beginning of an inhalation. Thus, we can measure the respiratory cycle length by extracting the duration between two consecutive peaks or valleys. In our thesis, we defined the valleys as the beginning of the respiratory cycle.

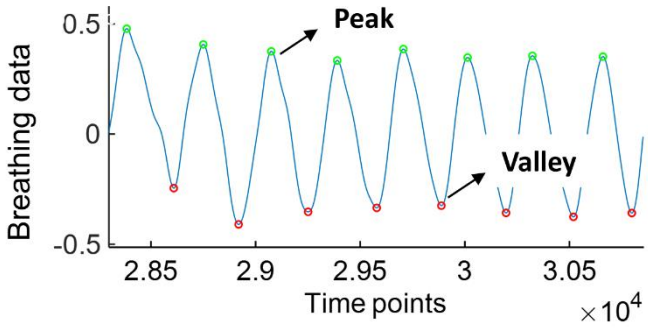


Figure 2.6 *– Example of a typical breathing data, where the red circles denote the valleys and the green circle denote the peaks.*

In (101), the author also noticed the presence of artifacts in some cycles, due to deep breaths or even due to participant movements. To remove the breathing cycles containing such artifacts, the author started by using the '*findpeaks*' function in Matlab. This function allows us to determine the aforementioned peaks and valleys. After finding two consecutive valleys or peaks, we can obtain the duration of the breathing cycle. Based on that, we can obtain the duration of each breathing cycle and subsequently exclude those with abnormal durations as outliers.

To do that, the author (101) computed the z-score using the *zscore* function in Matlab and removed the values with z-scores lower than -2.5. The cycles with z-scores higher than 2.5 were not removed, because it was believed that they were deep breaths and not artifacts. After that, he also studied the inhalation and exhalation length. The author (101) noticed that many artifacts were found in the middle of an inhalation or exhalation, making them erroneously defined as a peak or a valley, by the '*findpeaks*' function. This can happen at a point that was wrongly defined as a valley and the temporal distance from this point to the next valley is within the defined threshold. Therefore, this point, being within the threshold, is not considered an artifact. That said, the zscore function was also applied to the duration of inhalation and exhalation and the peaks and valleys that, in this approach, were defined as artifacts (z-score <2.5) were removed (the value was substituted by blank values - NaN).

2.7.3.1.1 *Examples of participants where respiration data had too many artifacts*

Backward masking study

Looking at the data of subjects 7 and 17, we noticed many tiny cycles, which makes it difficult to correctly identify the beginning and end moments of inspirations and exhalations. That said, we decided to exclude these participants from the analyses associated with respiration. Therefore, in the analyses involving respiratory data, we considered only 14 participants.

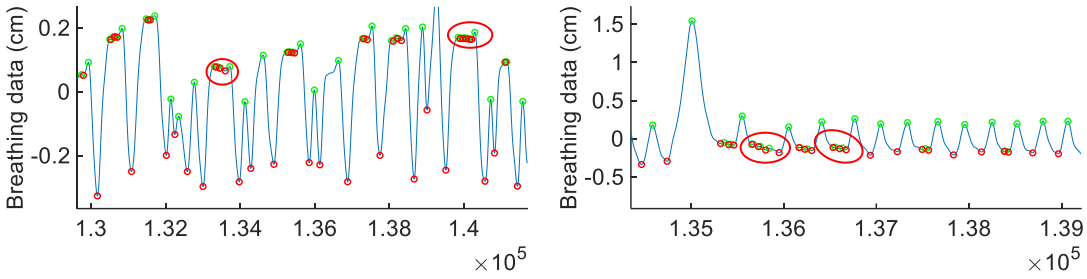


Figure 2.7 *- Breathing data from subject 7 and subject 17, respectively. In these two sub-figures we can visualize artifacts at several breathing cycles.*

Low contrast study

In subject 22, the previously mentioned behavior was also verified, therefore, this subject was discarded in the analyses associated with respiration.

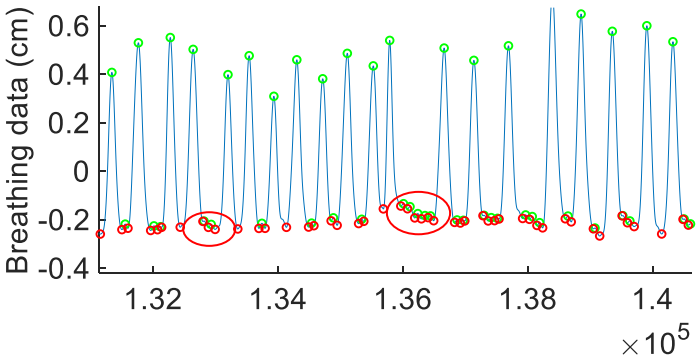


Figure 2.8 *- Breathing data from subject 22. In this figure we can visualize artifacts at several breathing cycles.*

2.7.3.2 Analysis of the respiratory activity

The respiratory signal is also a cyclic signal, so we can extract the cycle length, and, consequently, the respiratory rate. Contrary to what was done in the cardiac activity point, in this section we decided to only use the discrete approach.

To analyze how the cycle duration varied over time, we calculated the cycle duration of five cycles after the cue, inclusive, and one cycle before, and determined the cycles where the variation was significantly different from zero. To evaluate the respiratory cycle modulation just before stimulus onset and to study how its presentation influenced breathing rhythm, we studied the cycle duration variation one cycle before and one cycle after target onset. Then, we compared these variations across trial types (correct/incorrect trials in the backward masking study and recognized/unrecognized trials in the low contrast study) and evaluated the cycles where the difference was statistically significant.

Considering the discrete measure, we started by analyzing the cycle before and the cycle during the stimulus onset. However, if we consider the cycle at the time of the stimulus, its duration will already be influenced by the stimulus. Thus, we would not only be analyzing variations produced by the state of expectation, but also produced by the stimulus. Having said that, we decided to continue the analysis with just the duration of the cycle before stimulus onset. Based on that, we obtained the cycle duration before the stimulus is presented. Contrary to what was described in the ECG section, for respiration we decided to only study the cycle duration in an absolute and not in a relative way. We took this decision since the duration of the respiratory cycle is relatively long and sometimes the cue and the target were inserted in the same cycle, which prevented us from verifying the variations in the duration of the cycle in the state of expectation.

Repeating the same analysis performed for the cardiac cycle, we also compared the duration of the respiratory cycle with participant's accuracy (backward masking study) and stimulus detection (low contrast study). In addition to the aforementioned analyses, we also studied the correlation between the respiratory parameters and RT.

As mentioned, the respiratory signal is also a cyclic signal, so we also studied the phase of the cycle in which the stimulus was presented. As in the ECG, we perform a regression between the phase (sine and cosine) and the RT.

In this way, considering respiration, we included as features the cycle duration and phase in the classifier.

2.7.4 Eye movement and pupillography data

2.7.4.1 EyeLink data preprocessing

From the EyeLink we obtained two synchronized streams of data: eye-position samples and events, such as eye-movement events and subject responses or triggers (118).

To measure pupil size, the EyeLink extracts the area or the diameter of the pupil. First, the eye-tracker captures images of the eye using infrared illumination, which causes the pupil to contrast against the reflective background of the eye. To determine the pupil's boundaries, the EyeLink system used diverse algorithms. Then, based on the pixel dimensions, the system determined the diameter or the area of the pupil. Summarizing, for each point we had a positive value of pupil size, that represents, the area or the diameter. Allied to extracting pupil size, we extracted the gaze position (118).

The EyeLink software calculates the onset and end of the saccades and blinks in the data. In this way, we defined the latency of a blink or a saccade considering the onset of these events. This system records saccades using three thresholds, such as motion, velocity and acceleration. For instance, the velocity threshold is the eye-movement velocity that must be exceeded for a saccade to be detected. The EyeLink defines as begin and end of a blink the intervals where the pupil size is very small or the pupil in the camera image is missing or severely distorted by eyelid occlusion (118).

The data was initially acquired in *edf* format, so it was necessary to convert the files to *asc* format for analysis in Matlab. Missing data and blinks, as detected by the EyeLink software, were padded by 150 ms and linearly interpolated using an analysis script adapted from (80). Data were then normalized as percentage of the mean within each run i.e., for the pupil size, we subtracted the mean pupil size across all data points and subsequently divided by the mean.

Finally, in this step, we created a structure composed of nine channels. The first channel is composed of zeros and ones, where one denotes the time points in which data was interpolated. Then, the next five channels encompassed the signals acquired from the EyeLink. The seventh and eighth channels were composed of zeros and ones, with a value of one denoting the occurrence of a saccade or a blink. The last channel incorporated the pre-defined values referred to as triggers. So, in this step, we generated a structure with the data and with the triggers.

Finally, considering the time-windows used for the analysis, we removed the trials where the ET did not capture more than 50% of the data. Allied to this, we filtered the pupil size data with a bandpass filter limited between 0.1-6Hz, since the signal contained a lot of high frequency noise.

Outlier detection

The average spontaneous blink rate is said to be between 12 and 15/min (119). Regarding the data obtained by EyeLink, we found that participant 12 (backward masking study) had a very low rate of blinks. To do this analysis, we started by obtaining the rate of blinks per minute for each subject. Looking at the results, we saw that participant 12 blinked less than once per minute, and this low blinking rate did not allow for an accurate study of the effect of blinks. Thus, this subject was discarded in the analyses associated with blinking.

2.7.4.2 Analysis of the pupillary response

To evaluate the impact of the state of expectation in the pupil size, we studied the fluctuations in pupil size evoked by the cue within a time window ranging from 0.2s before and 10s after the cue onset (with baseline removed). We then identified the time points where these fluctuations were significantly different from zero. Subsequently, we studied these fluctuations just before visual stimulus onset and how the presentation of the stimulus influenced them. For this purpose, we analyzed the signal within a window of two seconds before and two seconds after the visual stimulus presentation. Furthermore, we compared the fluctuations in pupil size across trial types, using the same time windows. This comparison enabled us to establish the time points where statistically significant differences in the fluctuations were observed.

In order to investigate the effect of pupil-linked arousal on visual perception, we calculated the average pupil size, average pupil response (pupil variation in relation to pupil baseline measured in the 200 ms before cue onset), and average pupil derivative in an interval of 1s just before visual stimulus onset. This interval was chosen based on our own pilot studies. These pupil parameters were used to compare between correct/incorrect trials or recognized/unrecognized trials and to study the correlation between pupil parameters and reaction time. Then, the extracted pupil parameters were incorporated in the classifiers.

2.7.4.3 Analysis of blinking and saccadic activity

As described in the EyeLink data section, we obtained two signals composed of zeros and ones, with one being the start of a saccade or the start of a blink. With these signals, we were able to extract a blinks/saccades rate (averaged across all trials for each data point) for each participant.

To study the fluctuations in blinking/saccadic activity induced by the cue, we extracted the signal with a time window between 0.2s before and 10s after the cue onset (with baseline removed) and evaluated the time points where the variation in blinking/saccadic

activity was significantly different from zero. Additionally, we studied the signal just before and after stimulus onset by using an interval between 2s before to 2s after visual stimulus onset. Within this context, we evaluated whether the fluctuations in blinking and saccadic patterns induced by the cue were statistically different from zero. Lastly, we compared the variations in blinking/saccadic activity across trial types.

Subsequently, we conducted an analysis of blinking and saccadic activity by contrasting the activity parameters between correct/incorrect trials or recognized/unrecognized trials. Furthermore, we explored the potential correlations between these parameters and reaction time. These parameters were later used as features in the classifiers.

For feature extraction related to the rate of saccades and blinks, to maximize the probability of detecting blinks and saccades that might affect the stimulus processing, we decided to use as an interval of interest the interval between 1s after the cue and the target onset. We decided to not include the first second after the cue, because it is known that when we get an alarm signal or any abrupt or unexpected stimulus, our body automatically reacts to protect the eye, which may cause a blink, and to look for the possible threat, which may cause a saccade, and this reaction does not reflect changes produced by this state of alertness. We used the average saccades and blink rate in this interval and compared across trial types and used them as features in the classifiers.

Moreover, we expect that the closer a saccade is to the stimulus display, the greater the probability of the participant to miss the stimulus. Based on this, we decided to include a further feature where we estimated the temporal distance between the last saccade and the stimulus onset. We defined the maximum temporal distance as the temporal distance between one second after cue onset and stimulus onset, as described above. Thus, in the trials in which no saccade or blink occurred in the considered interval, we considered the temporal distance as the limit.

2.8 Classification models

As mentioned before, the aim of our study was to examine the influence of the pre-stimulus state on visual processing. To do so, we created a first classifier that used the pre-stimulus neural and physiological signals (pre-stimulus being the activity before the appearance of the visual target) as input to classify visual performance (visual performance model). In addition, we used, in a second classifier, the post-stimulus activity (neural response evoked by the visual targets) to classify the category of the images presented and investigated the impact of the pre-stimulus activity on target representation (image category model). In the visual performance model, we studied the ability of the classifier to predict trial accuracy, in the backward masking study, and trial visual recognition, in the low contrast study. Our goal with this model was to investigate whether the pre-stimulus activity influences the detection of the stimulus regardless of the category (the pre-stimulus activity modulates the ability of the participants to recognize the presence of an object - visual sensitivity). In the image category model, in turn, we investigated if the pre-stimulus activity modulated how the stimulus was decoded.

We used two different types of classifiers, the SVM and the CNN. The SVM was used to test the ability of the physiological signals to classify visual perception. The CNN (EEGNet) was used when EEG signals were processed as input features.

2.8.1 Data

We started our analyses regarding classifiers, by applying the classification model individually to each subject. However, one of the problems we faced was the fact that, for each participant, we had relatively few trials, which would influence the performance of the classifier. To solve this problem, we decided to make a model that receives as input the trials of all participants to enhance the classifier's ability to recognize patterns and consequently achieve better performance. In this new approach, we normalized each participant's input data, with the aim of preserving the integrity and variability of characteristics within participants.

In the backward masking study, we verified that the participants achieved an average accuracy exceeding 75% (results section). Thus, we observed a higher number of correct trials compared to incorrect trials. In the low contrast study, as described, we focused on studying the recognition of the trial. As observed in the results section, participants, on average, recognized more than 50% of the trials, resulting in a higher number of recognized trials than the unrecognized ones. Consequently, for both studies we had imbalanced data. To address this problem, we used the class weights technique. This technique involves

assigning higher weights to the minority class, enabling the model to prioritize and assign more importance to its samples during training and reduce bias towards the majority class.

2.8.2 Training and test sets

As mentioned in the previous subsection, we started our analyses by applying the classification models to each participant individually. However, each participant had relatively few trials. To combat this problem, we decided to use an approach that made use of all trials of the participant to train and test the algorithm, the cross-validation approach. This approach consisted of dividing the trials into train and test, which was repeated *k* times. To our classifier, we choose a value of *k* equal to five. That is, the network was run five times and, consequently, five performances were obtained. Thus, the final performance of the network was the average of the five performances. Then the data for training was divided again, where 80% was for training the network and 20% for validating.

When combining the data of all participants, we implemented a strategy where 70% of each participant's trials were extracted for training purposes, while the remaining 30% were reserved for testing, to ensure the inclusion of data from all participants in the test set. Thus, the overall distribution of trials resulted in 70% for training data and 30% for test data. To enhance the reliability of our findings, we opted to execute the classifier multiple times. Specifically, with the SVM algorithm, we ran 20 times. However, due to the significant computational requirements associated with CNN, we limited the number of runs to 15.

2.8.3 Performance metrics

We used three distinct metrics to assess the classifiers' performance: the AUC, accuracy, and balanced accuracy for when the classes were not balanced. In the particular case where the classes are not balanced (when classifying visual performance), the AUC is particularly useful, as it assesses the model's capability to rank samples correctly rather than considering a single threshold for classification. In order to standardize the analyses presented in the thesis, we decided to use the AUC as the performance metric for the remaining analyses, whenever possible.

For conciseness, in the body of the thesis, only the results associated with the AUC metric are presented. The results that evaluate performance through accuracy and balanced accuracy can be found in the appendix.

2.8.4 Algorithms

2.8.4.1 Support vector machine's algorithm structure

The data that we obtained from the physiological signals resulted from a discrete measure applied to the pre-defined pre-stimulus interval, that is, for each measure we had a value associated with a particular trial. Given the relatively small dimension of the dataset and its complexity, due to its non-linearity, we opted to employ the SVM algorithm instead of the CNN that was optimized for EEG.

The SVM is an algorithm that is able to, as mentioned, perform both linear and non-linear classification. Considering that the biological signals contain non-linear dynamical characteristics, we opted to employ a non-linear model (120). For that, it was necessary to choose the kernel trick that best applies to our data type. We ultimately decided to choose the RBF kernel, as it is more robust and able to capture the more complex relationships between data.

- **Optimization** - The optimization of SVM hyperparameters is a crucial point in the development of a SVM algorithm, since it allows a better performance and generalization ability of the model. We decided to optimize the sigma σ and the penalty c parameters with the '*OptimizedHyperparamenter*' option provided by Matlab. This option makes use of the Bayesian optimization, that involves constructing a probabilistic model. To evaluate this model, it is necessary to use an acquisition function. We decided to use the default acquisition function provided by MATLAB: *'expected-improvement-per-second-plus'* (89,121). The expected improvement function measures the probability that a particular candidate sample is better than the current best sample (90,121). One way to improve the performance is to optimize the acquisition function. So, we used '*expected improvement per second plus'*, because this approach prioritizes the acquisition of points that are not only likely to be good, but that are also likely to be evaluated quickly (90,121).

2.8.4.2 Convolutional neural network's algorithm structure

To effectively capture the characteristics of the signal across all EEG data points and channels within the predefined time interval, it was imperative to select an algorithm capable of processing a very large amount of data. In addition, the chosen classifier needed to be able to capture both spatial and temporal patterns in a time series format. In this way, we opted for the CNN algorithm due to its proficiency in extracting spatial and temporal patterns through the application of trainable filters and assigning importance to these patterns using trainable

weights. For these analyses we decided to use the EEGNet algorithm due to its architecture that was specifically designed and optimized for processing EEG data (96).

In the classification analyses, where we used only EEG data as input, the EEGNet algorithm was used. However, as previously indicated, the aim of our project was to investigate the influence of pre-stimulus activity on visual perception. To this end, we used as features not only the continuous EEG signal, but also the other measures of the different physiological signals. However, the EEGNet algorithm was not designed to incorporate additional data beyond the continuous EEG signal. So, for these, we adapted the EEGNet classifier, which we called *multimodal classifier*. This algorithm allows the concatenation of the various measures to the continuous EEG signal.

Similar to the EEGNet algorithm, in the multimodal algorithm, spatial and temporal features are extracted from the continuous EEG signal. Thus, the multimodal algorithm also comprises the two aforementioned blocks described above (block one for learning how to filter in time and in space the input EEG and block two for learning how to resume in time the information). Additionally, the measures obtained from the various physiological signals are introduced into the network in the one before the last layer. This is achieved by concatenating the features obtained through processing the EEG signal with the aforementioned measures. Lastly, the resulting features go through a classification block (Figure 2.9).

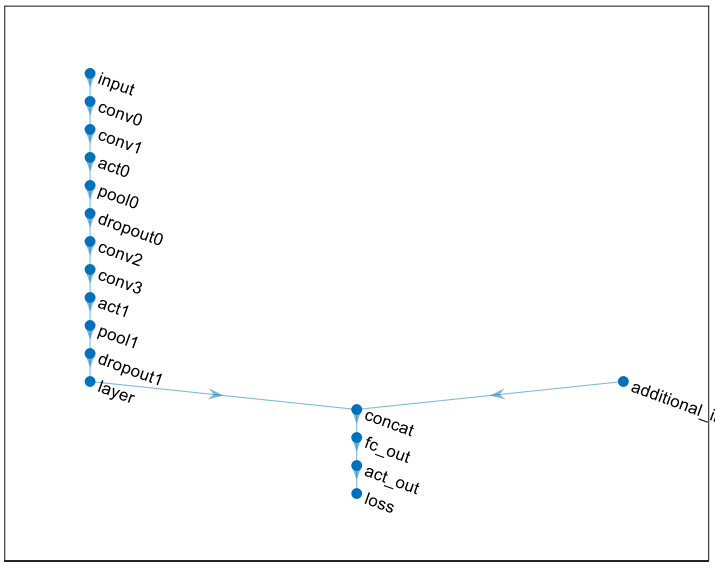


Figure 2.9 *– Overall visualization of the multimodal classifier's architecture.*

The process of designing and optimizing the structure of the model was one of the first analyses carried out. Considering that we used an algorithm previously developed in (96), the basis of the structures of the classifier had already been predefined. That said, for this classifier it was only necessary to define the parameters (table 6) and optimize some of them.

- **Optimization** - In order to facilitate and enhance the efficiency of the optimization process, we decided to focus only on optimizing three specific parameters: mini-batch size, probability dropout and temporal filters. Selecting the appropriate mini-batch size allows us to achieve training stability and improve generalization performance of the model. The optimization of the temporal filters allows us to choose the best value that facilitates the extraction of all temporal and spatial characteristics. Lastly, we optimized dropout probability to prevent overfitting and improve the generalization of the algorithm. To choose the best parameters for the EEGNet classifier, we analyzed the performance of the EEGNet classifier using ERP data from the backward masking study and evaluated its ability to classify the images category. We ran the classifier five times for each participant using the 5-fold cross validation method and compared the classification AUC for each set of parameters. The combination of parameters that offered the best performance was mini-batch size - 64; temporal filters – 20; probability dropout 0.5. The results of these tests can be found in Figure 2.10.

Solver – model optimization	*'adam'*
Learning rate	*0.01*
MaxEpochs	*200*
L2Regularization	*0.001*
Mini-batch size	*64*
Shuffle	*'every-epoch'*
ValidationFrequency	*floor(length(y_train)/mini_bs)*

Table 6 *- Training options of the CNN algorithms.*

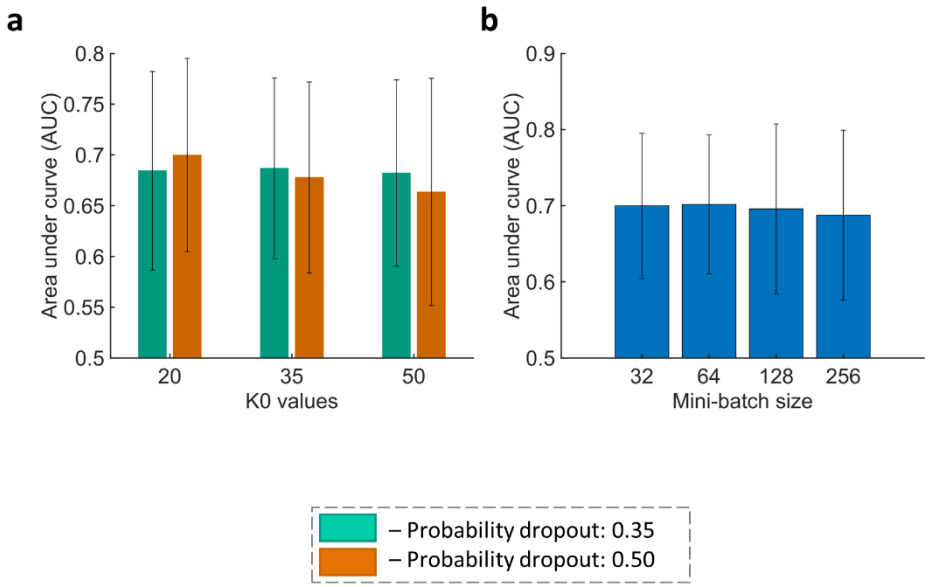


Figure 2.10 *– Optimization of the hyper-parameters.* ***a*** *Performance of the EEGNet using a combination of various values of temporal filters and various values of dropout probability.* ***b*** *Performance of the EEGNet using different values of mini-batch size. In sub-figure* ***b****, the temporal filter and the dropout probability were already optimized. The green bars represent the combination of different temporal filters with a dropout probability of 0.35. In turn, the orange bars represent the combination of different temporal filters with a dropout probability of 0.50.*

2.8.5 Analyses performed

2.8.5.1 Visual performance model

In this model, our goal was to investigate the potential impact of pre-stimulus brain and body activity on stimulus detection, i.e., the influence on visual performance.

To study the influence of pre-stimulus body activity, we used an SVM algorithm that receives as input a vector with the parameters of a certain physiological signal in each trial. We evaluated the ability of the classifier to predict visual performance by combining all parameters or by using each parameter individual. Our aim when studying each parameter individually was to evaluate the contribution of each parameter to the classifier. In other words, the objective was to understand which individual component of the cardiac activity, for instance, exerted a more substantial impact on the classifier's performance, which could potentially represent a heightened influence of that component on stimulus detection. It is essential to highlight that this was one of the last analyses to be carried out and, due to lack of time, it was exclusively performed for the all-trials approach.

We then used the EEGNet algorithm to investigate the potential influence of brain activity on visual performance. For that, we used the EEG signal as input. One of the

challenges regarding this analysis was selecting the optimal pre-stimulus interval. Considering that the minimum interval of state of expectation was two seconds, we decided to choose this interval as the maximum interval of interest. Allied to this, we also choose as a possible interval of interest the one second interval and the 0.5 seconds interval. It is important to highlight that initially, we conducted the analyses for each subject individually. Comparing the performance obtained for the various intervals, we can see that, for both the backward masking study and the low contrast study, the only interval that provides a performance statistically superior to 0.5 is the one second interval in the low contrast study (Figure 2.11 and Table 7). Consequently, to ensure standardization and facilitate a meaningful comparison between the two studies, we made the decision to choose a one-second pre-stimulus for both studies.

	Backward masking study	**Low contrast study**
Time of pre-stimulus (s)	**One-sample *t*-test**	**One-sample *t*-test**
0.5	$t(15) = -0.471$, $p = 0.645$	$t(15) = 0.128$, $p = 0.900$
1	$t(15) = -0.333$, $p = 0.744$	$\mathbf{t(15) = -2.487, p = 0.023}$
2	$t(15) = -0.636$, $p = 0.535$	$t(15) = -1.267$, $p = 0.221$

Table 7 *– Statistical analyses of the classifier considering different intervals of pre-stimulus activity. We used one-sample t-test to evaluate whether the classifier had a behavior statistically superior to 0.5. The statistically significant analyses are highlighted.*

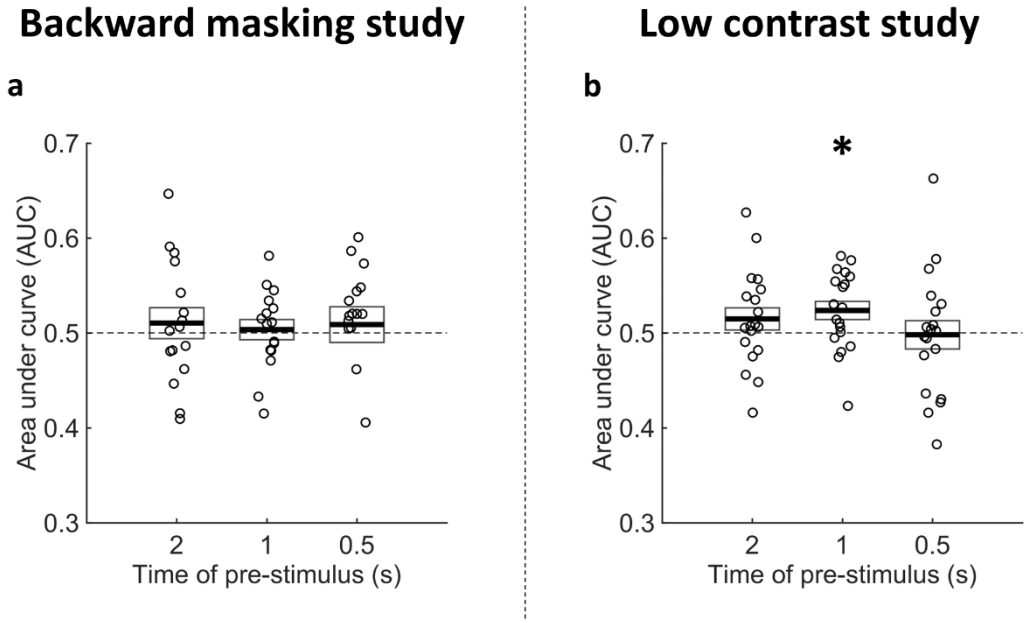


Figure 2.11 *– In the backward masking study, none of the selected intervals demonstrated a statistically significant performance higher than 0.5. In turn, in the low contrast study, when using a one-second interval, the performance was statistically greater than 0.5.* ***a*** *AUC using the EEGNet model to study the influence of different EEG pre-stimulus intervals (backward masking study).* ***b*** *AUC using the EEGNet model to study the influence of different EEG pre-stimulus intervals (low contrast study). Each bar shows the results for the respective pre-stimulus interval. The black line is the average performance and the*

*rectangle represents ± standard error of the mean. Each circle within the representation denotes each participant. *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

Lastly, we incorporated the body pre-stimulus activity to find out if it contained supplementary information that could be recognized by the classifier, enhancing its performance. In summary, in the first analyses the network was only fed with the continuous EEG signal, while in the second analyses the network was fed with this signal and with the measures described in the previous section.

In order to ascertain the capacity of the multimodal classifier to extract activity patterns from physiological signals, we implemented a control algorithm. This algorithm uses the multimodal classifier while setting the EEG to zero, with the primary objective of determining the classifier's capability to predict visual performance based only on pre-stimulus body activity. This approach was chosen to ensure that any potential absence of significant improvements in the classifier's performance, when combined with the pre-stimulus body activity, is attributed to the ineffective contribution of this activity to the classifier and not because the classifier is not able to extract activity patterns from these signals.

2.8.5.2 Image category model

In this model, we started by assessing the classifier's capacity of, through the EEG signal evoked by the visual stimuli (event-related potentials (ERPs)), discriminating the stimulus category.

After studying whether brain and body pre-stimulus activity influence the detection of the stimulus and, consequently, the visual performance, we studied if this activity modulates the neural representation of the visual stimulus. With this aim, we combined the body and brain pre-stimulus activity with the ERPs and evaluated if this combination enhanced the classifier's performance.

2.9 Statistical analyses

In all the statistical analyses, a cut-off of $p < 0.05$ was used to define significance. All statistical tests were run in Matlab.

2.9.1 Data analysis

For each variable used, we first evaluated if the data distribution was normal using the Shapiro-Wilk test implemented in Matlab's *swtest* function. This test tests the null hypothesis that a given sample follows a normal distribution. Thus, if the p-value is lower than a chosen significance level, the null hypothesis is rejected, which suggests that the data do not follow a normal distribution. If the variables followed a normal distribution, parametric statistical tests were used (one-sample t-test, paired t-test, and Pearson correlation). If the variables did not follow a normal distribution, non-parametric statistical tests were used (Wilcoxon test - the *signrank* function of Matlab and Spearman's correlation). Thus, to assess whether two variables were statistically different, for example, we first studied their distribution and, based on that, we used the t-test or the Wilcoxon test. In turn, if we wanted to assess a possible correlation between two variables, we applied the normality tests again and, subsequently, the correlation tests, Pearson correlation or Spearman's correlation.

To analyze the correlation between the phase of the cardiac cycle and the respiratory cycle at which the stimulus was presented and the RT, we used a slightly different approach. In contrast to the previous correlations, when studying the phase, we considered two variables: the sine and cosine and employed linear regression models, to establish the relationship between the mentioned variables, the sine and cosine of the phase, and the RT. Considering that the phase corresponds to a sine and cosine transformation, we can use it as circular predictor of RT in a regression model with the coefficients β_1 and β_2 (122). The regression model can be represented by the following expression:

$$RT_i = \beta_0 + \beta_1 \cos\theta_i + \beta_2 \sin\theta_i + \epsilon$$

Equation 9 *– Regression model expression.*

where RT_i is the RT on trial i, θ_i is the phase at which the stimulus is presented in trial i, β_0 is the intercept term and ϵ the error term. To obtain the coefficients, we used the Matlab's function *regress*. To determine whether the coefficients are significantly different from zero, we started by examining whether the coefficients followed a normal distribution using the

swtest function in Matlab. Then, based on the distribution of the coefficients, we utilized either the one-sample *t*-test or the Wilcoxon test.

To study the variations in body and brain activity evoked by the cue and to compare across trials type, we used the *mult_comp_perm_t1* function developed by David Groppe (123). This function analyses all data points and corrects for multiple comparisons using the permutation method to compute the *p*-values. The permutation tests establish the likelihood that an observed behavior arose "by chance". This test iteratively resamples the observed data to determine the *p*-value for the given test (124). With the same function (123), we examined the differences observed in ERPs between the car and house trials. The objective with this analysis was to find out whether the two categories produced significant variations in the neural processing of the stimulus and subsequent decision-making processes.

2.9.2 Statistical analyses of the classifier outputs

To study the ability of the classifier to perform better than the chance level, we first employed a one-sample *t*-test. When using this test, our objective was to study whether the performance statistically exceeds 0.5 when using AUC as the performance metric or surpasses 50% when accuracy was chosen as the performance metric.

However, the doubt arose whether each set of classifiers was effectively better than the chance level or whether the result was related to the fact that the data were not balanced (especially in the visual performance model). In this way, this outcome does not ensure that the classifier's behavior is non-random. Consequently, we made the decision to replace one-sample *t*-test for an alternative statistical test, namely the permutation test, to determine whether each classifier individually exhibited a behavior distinct from randomness. In this way, we started by extracting the performance of each classifier using the true labels. Next, we randomized the true labels and, again, extracted the performance of the classifier (both AUC and accuracy). This process was repeated 1000 times. To find out whether the classifier exhibited superior behavior compared to random chance, we compared the performance of the classifier with the true labels to the performance with the randomized labels. We considered that the classifier had a superior behavior to the random one if its performance exceeded the performance with the randomized labels in at least 95% of the iterations (*p*-value = 0.05). However, as mentioned, each classifier was run multiple times. That said, for a given analysis, this permutation test was applied to each individual classifier. Thus, it was considered that, for example, with cardiac activity, we could predict the recognition of the stimulus if at least 50% of the classifiers had a behavior superior to random chance.

It should be noted that, in all the statistical analyses of the classifier outputs, we determined whether the classifier exhibited an ability of distinguishing in a non-chance way.

It is essential to highlight that the inclusion of permutation tests occurred exclusively when the approach of incorporating all subjects in the classifier was adopted. In summary, when the analyses were carried out within the subject, we employed the one-sample t-test.

2.9.2.1 Visual performance model

We started by studying the capacity of the classifier to distinguish visual performance using brain and body pre-stimulus activity.

We began by studying the influence of each body physiological signal with the SVM classifier and by evaluating the individual contribution of different parameters within each signal. In this way, we performed a paired-sample *t*-test to assess the difference between the performance obtained with each measure and the one obtained with the combination of all measures of a certain signal (all ECG measures, for example).

Then, we studied how pre-stimulus body activity improved the performance of the multimodal classifier. We started by evaluating the capacity of the multimodal classifier to extract activity patterns from the pre-stimulus body activity. We studied each physiological signal individually and then combined all physiological signals. To ascertain the contribution of each physiological signal when using all the signals combined, we used a paired-sample *t*-test.

Subsequently, we evaluated the classifier's ability to predict visual performance using only the EEG pre-stimulus activity. Then, we included each measure, with the expectation that each measure would improve the performance of the classifier. To assess the improvement of the performance when incorporating each physiological signal, we compared the classifier's performance using only the pre-stimulus neural activity to the performance when each physiological signal was included. To determine whether the observed differences were statistically significant, we performed a paired-sample *t*-test.

2.9.2.2 Image category model

Considering that our initial analyses were performed within participants, the following analyses were done within each individual participant. The first analyses focused on analyzing the ability of the classifier to discriminate the category using only the ERPs. Then, we compared the classifier's performance with the actual performance of the participant in the task (the hits and misses of the participant). To facilitate the comparison of the performances, we opted to use only the accuracy, instead of also using the AUC metric. To study if there is a statistically correlation between both performances, we ran the Matlab's function *corrcoe*f.

After verifying that we were effectively able to discriminate the trial category using only ERPs within participant analyses, the same analysis was performed considering data from all participants.

The next step involved integrating the pre-stimulus activity into the classifier and finding out if it modulated the stimulus decoding. To assess the improvement of the performance when incorporating each physiological signal, we compared the classifier's performance when only considering the post-stimulus neural activity to its performance when incorporating each signal individually and when incorporating all signals combined. To determine whether the observed differences were statistically different, we performed a paired-sample *t*-test.

3 Results and Discussion

In this chapter, we will present and discuss the results of the analyses we performed, using the methods outlined in the previous section.

3.1 Task performance

In the backward masking study, the average accuracy percentage across all participants was 80.1%, with a standard deviation of 9.7% (Figure 3.1 - a)). During data acquisition, participants reported that they detected more cars than houses. In fact, the mean accuracy for houses was 71.5%, with a standard deviation of 17.5%, whereas the mean accuracy for cars was 90.2%, with a standard deviation of 9.7% (Figure 3.1 – b)). Applying statistical tests, we can conclude that the performance in car trials is significantly higher than the performance in house trials (Shapiro-wilk test: W =0.826, p = 0.006; Wilcoxon test: Z = -3.154, p = 0.002).

In the study that we used low contrast images, participants recognized 65 ± 13% [mean ± standard deviation (SD)] of the trials that contained images (Figure 3.1 - c)). Initially, we expected a percentage of recognition closer to 50%; however, we observed a higher recognition. We believe that this result is related to the fact that the adaptive threshold procedure did not include a warning cue and therefore did not induce the same state of attention as in the task and also there could have occurred some perceptual learning that improved performance over time. In turn, participants only recognized 8% of the scrambled images (with a standard deviation of 9.65%) (Figure 3.1 - c)). This low percentage indicates that the participants were following task instructions.

Participants were instructed to report that they did not recognize the stimulus when they did not see any object. Therefore, we expected that the accuracy would be 50% on the unrecognized trials, due to the random choice. Our results showed that in unrecognized trials, the participants correctly answer 57.71% of the trials (with a standard deviation of 8.14%). In turn, in recognized trials, the participants correctly answer 91.21% (with a standard deviation of 6.82%) of the trials (Figure 3.1 - d)). As expected, the accuracy values in recognized and unrecognized trials were significantly different (Shapiro-wilk test: W = 0.984, p = 0.980; paired-sample t-test: $t(18)$ = -18.091, p < 0.001). These findings suggest again that participants were following task instructions.

Finally, an analysis was carried out to evaluate the difference in car trials recognition and house trials recognition. In the backward masking study, we can see significant

differences between both categories, which suggests that participants, in this study, continued to detect more cars compared to houses (Shapiro-wilk test: $W = 0.9514$, $p = 0.4166$; paired-sample t-test: $t(18) = -4.391$, $p < 0.001$; Figure 3.1 – c)). This result might be attributed to the fact that a car has salient characteristics, such as wheels or a bumper, that are easier to identify than those of a house.

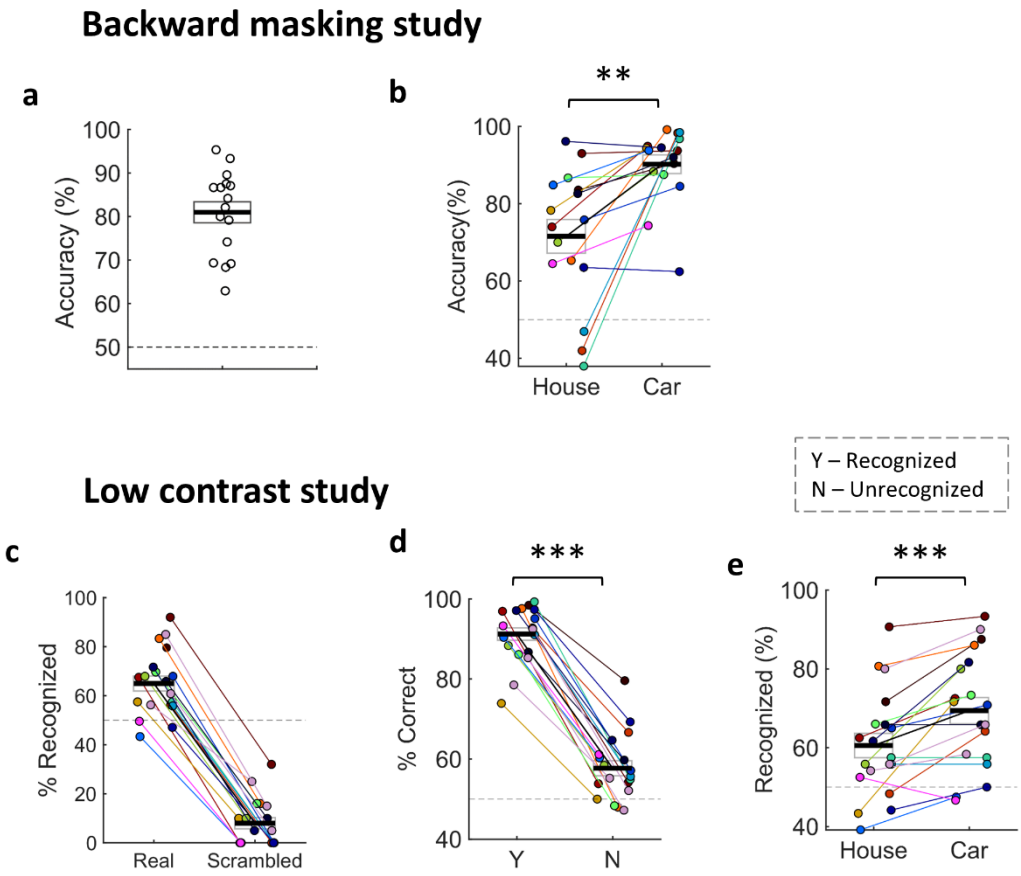


Figure 3.1 *– Behavioral results.* ***a*** *Percentage of trials correctly answered for each participant (backward masking study).* ***b*** *Accuracy for house and car trials study (backward masking study).* ***c*** *Percentage of trials reported as recognized for real and scrambled images (low contrast study).* ***d*** *Accuracy in recognized and unrecognized trials of real images (low contrast study).* ***e*** *Percentage of recognized trials in both categories of our study (low contrast study). The black horizontal line represents the mean across participants and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

The RT is a measure of how quickly an organism reacts to a stimulus. In this way, RT can be used as a measure of response confidence and will reflect the difficulty in detecting the stimulus (125). As expected, in the trials where participants correctly identified the stimulus category, the RT was significantly shorter compared to the trials where participants misidentified the stimulus category (backward masking study; Figure 3.2 – a); Shapiro-wilk test: $W = 0.858$, $p = 0.018$; Wilcoxon test: $Z = -3.5162$, $p < 0.001$). These findings are in line with previous studies that show a decrease in RT related to an increase in accuracy (125).

The procedure was repeated in the low contrast study. However, considering that, in this version, we were able to study stimulus detection, instead of studying the relation between RT and accuracy, we decided to relate the RT with the recognition of the trial. As anticipated, trials where participants recognized the stimulus showed significantly shorter RTs in comparison to when they did not recognize it Figure 3.2– b); Shapiro-wilk test: W = 0.726, p <0.001; Wilcoxon test: Z = -3.783, p < 0.001).

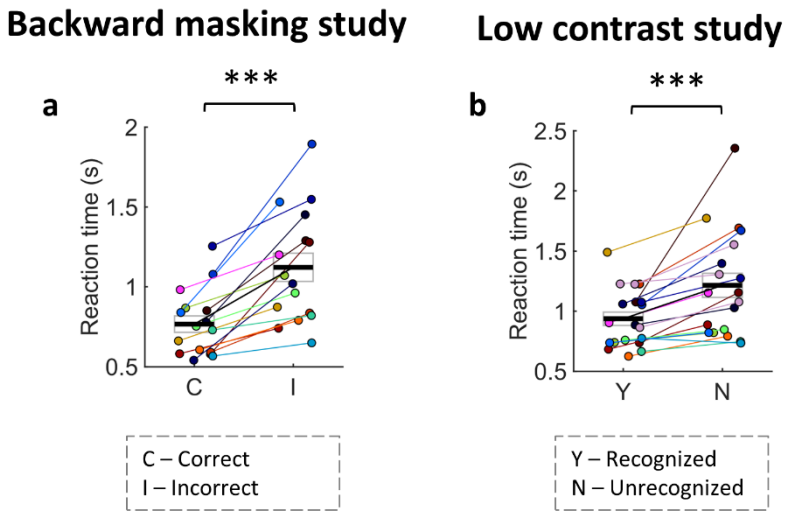


Figure 3.2 *– Reaction time is lower in recognized trials, as well as in correct trials.* ***a*** *Relation between RT and task accuracy (backward masking study).* ***b*** *Relation between RT and stimulus recognition (low contrast study). The black horizontal line represents the mean across participants and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2 Effect of pre-stimulus neural and physiological activity on visual perception

We started by evaluating the changes in body and brain activity induced by the state of expectation. Subsequently, we investigated the association between pre-stimulus activity and visual perception. Then, we used an SVM model to classify visual perception (correct vs incorrect trials in the backward masking study and recognized vs unrecognized trials in the low contrast study) using as features the described parameters of the different physiological signals. In turn, when using the EEG pre-stimulus activity to classify visual perception, we decided to use a CNN model.

3.2.1 Effect of pre-stimulus cardiac activity on visual perception

3.2.1.1 Study of the time course of heart rate variation evoked by the warning cue

Studies developed in the past years suggest that the state of expectation activates, among other aspects, the parasympathetic system, which causes a decrease in heart rate (82). These findings are supported by the results obtained in both studies. In both Figure 3.3 – a) and Figure 3.3 – b), we can see that the auditory cue evoked a cardiac deceleration. This deceleration is followed by a cardiac acceleration, after visual stimulus onset, as we can see in Figure 3.3 – c) and Figure 3.3 - d).

In the backward masking study, when comparing correct with incorrect trials, we see a significant difference in cardiac deceleration with correct trials presenting a stronger decrease in heart rate (Figure 3.3 - e)). We do not see any difference when comparing recognized vs unrecognized trials in the low contrast study. Also, just before target onset (Figure 3.3 - g) and Figure 3.3 - h)), we can see that cardiac deceleration is stronger in correct (backward masking study) and recognized (low contrast study) trials. However, this difference was not statistically significant. After target presentation, we see that, in correct trials (backward masking study) and in the recognized trials (low contrast study) the heart starts to accelerate around 500 ms which corresponds to the moment of the decision, i.e., the heart starts accelerating around the time of the motor response. Considering that, in incorrect/unrecognized trials, the subject takes longer to make the decision, so, we expect the heart to start accelerating later. These conclusions can be verified in Figure 3.3 – g) and Figure 3.3- h).

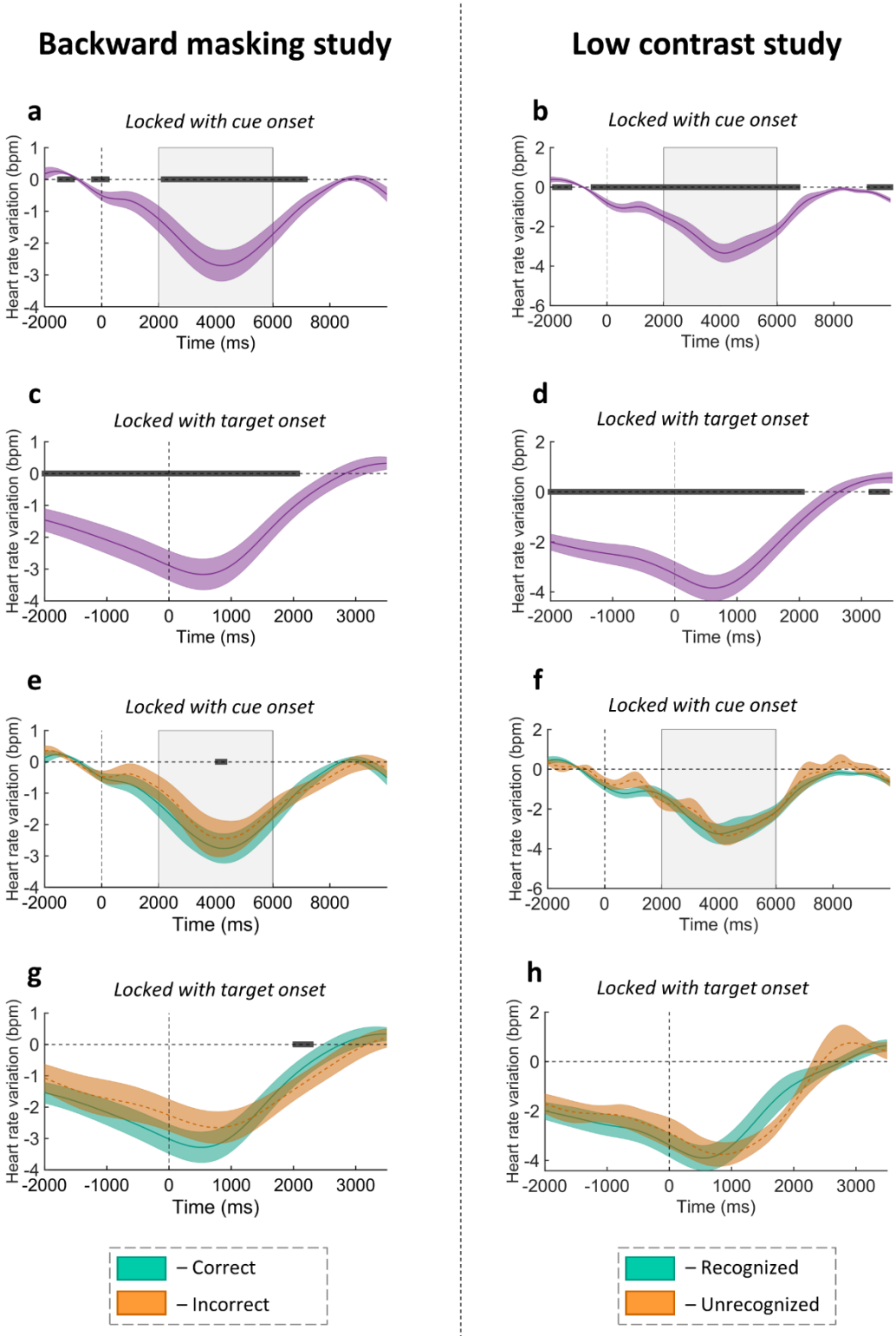


Figure 3.3 *– The cue induced cardiac deceleration.* ***a and b*** *Cue-locked heart rate modulation (a - backward masking study; b - low contrast study).* ***c and d*** *Target-locked heart rate modulation (c - backward masking study; d - low contrast study). The gray horizontal line represents the time points where the cardiac response is significantly different from zero (p < 0.05). The gray rectangle represents the time window in which the stimulus is presented.* ***e*** *Cue-locked heart rate modulation in correct trials (green curve) and incorrect trials (orange curve) (backward masking study).* ***f*** *Cue-locked heart rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study).* ***g*** *Target-locked heart rate modulation in correct trials (green curve) and incorrect trials (orange curve) (backward masking study).* ***h***

*Target-locked heart rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study). In **e**, **f**, **g**, and **h**, the gray horizontal line represents the significant time points where the cardiac response is significantly different across the two conditions ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. In all graphs, data are represented as mean ± standard error of the mean across participants.*

3.2.1.2 Analyses of cardiac features used in the classifiers

For the analyses of the effect of pre-stimulus cardiac activity in visual perception using classifiers, we extracted three different types of features: heart rate, heart rate variation and cardiac phase. First, we studied how these were related to task performance in both studies and then we investigated if an SVM classifier was able to use them to predict the participant's performance.

3.2.1.2.1 Absolute heart rate and heart rate variation

We started by analyzing the absolute heart rate. In the backward masking study, as illustrated in Figure 3.4 – a), we observe that the heart rate for correct and incorrect trials is practically the same, which suggests that the participant's accuracy is not significantly associated with the absolute heart rate (Shapiro-wilk test: $W = 0.981$, $p = 0.974$ and paired-sample t-test: $t(15) = 0.743$, $p = 0.469$). In the low contrast study, in turn, in Figure 3.4 – c), we can see that in unrecognized trials the heart rate is significantly higher (Shapiro-wilk test: $W = 0.814$, $p = 0.003$ and Wilcoxon test: $Z = -1.972$, $p = 0.040$).

Initially, when we explored the effect of heart rate in the backward masking study, since we did not see any correlation between heart rate and task accuracy, we hypothesized that the cardiac cycle had no effect on visual processing. However, in the low contrast study, we can see that there are significant differences in heart rate in the analyzed conditions. The results, in this study, are in line with those found in (2), which suggest that the duration of the cardiac cycle at the moment of stimulus presentation influences the visual processing. Thus, we can consider that the results of the backward masking study were affected by the use of the mask and by the fact that several trials were answered correctly by chance, despite the physiological state not favoring such correct responses.

Moreover, we investigated the relationship between RT and heart rate. We are going to start with the backward masking study. Since in absolute heart rate there was no significant difference between correct and incorrect trials, we expected that a correlation between RT and heart rate would not be observed. In Figure 3.4 – b), the correlation values are not statistically different from zero, which goes accordingly to what we were expecting (Shapiro-wilk test: $W = 0.993$, $p = 0.118$ and one-sample t-test: $t(15) = -0.560$, $p = 0.584$). In the low contrast study, on the contrary, we observed a higher heart rate in unrecognized trials.

Based on that, we anticipated a positive correlation between heart rate and RT. We expected this, because, when we do not recognize the stimulus, we tend to take longer to decide which option to choose. In Figure 3.4 – d), we can see the precise correlation that we had anticipated (Shapiro-wilk test: $W = 0.962$, $p = 0.602$ and one-sample t-test: $t(18) = 2.850$, $p = 0.011$). The results, in the low contrast study, suggest that the heart rate influences the RT, i.e., the heart rate has a role in visual processing. Again, in the backward masking study, we do not see this influence, probably, due to the influence of the mask.

After studying the influence of the absolute heart rate, we studied the heart rate modulation. In both studies, there is no statistically significant difference between heart rate variation and participant's accuracy ((Figure 3.4 – e) and (Figure 3.4 – g)) (backward masking study: Shapiro-wilk test: $W = 0.955$, $p = 0.495$ and one-sample t-test: $t(15) = -0.981$, $p = 0.342$; low contrast study: Shapiro-wilk test: $W = 0.876$, $p = 0.020$ and Wilcoxon test: $Z = -0.926$, $p = 0.655$).

Similarly to the analysis conducted for heart rate, we also examined the relationship between heart rate variation and RT. As for the absolute heart rate, the correlation between heart rate variation and reaction time is not significant in the backward masking study (Shapiro-wilk test: $W = 0.9068$, $p = 0.094$ and one-sample t-test: $t(15) = 0.338$, $p = 0.740$; Figure 3.4 – f)) but is significant in the low contrast study (Shapiro-wilk test: $W = 0.974$, $p = 0.762$ and one-sample t-test: $t(18) = 2.126$, $p = 0.042$; Figure 3.4 – h)). The results obtained in the low contrast study suggest that the heart rate variation can influence the time needed to make the decision and probably the visual processing.

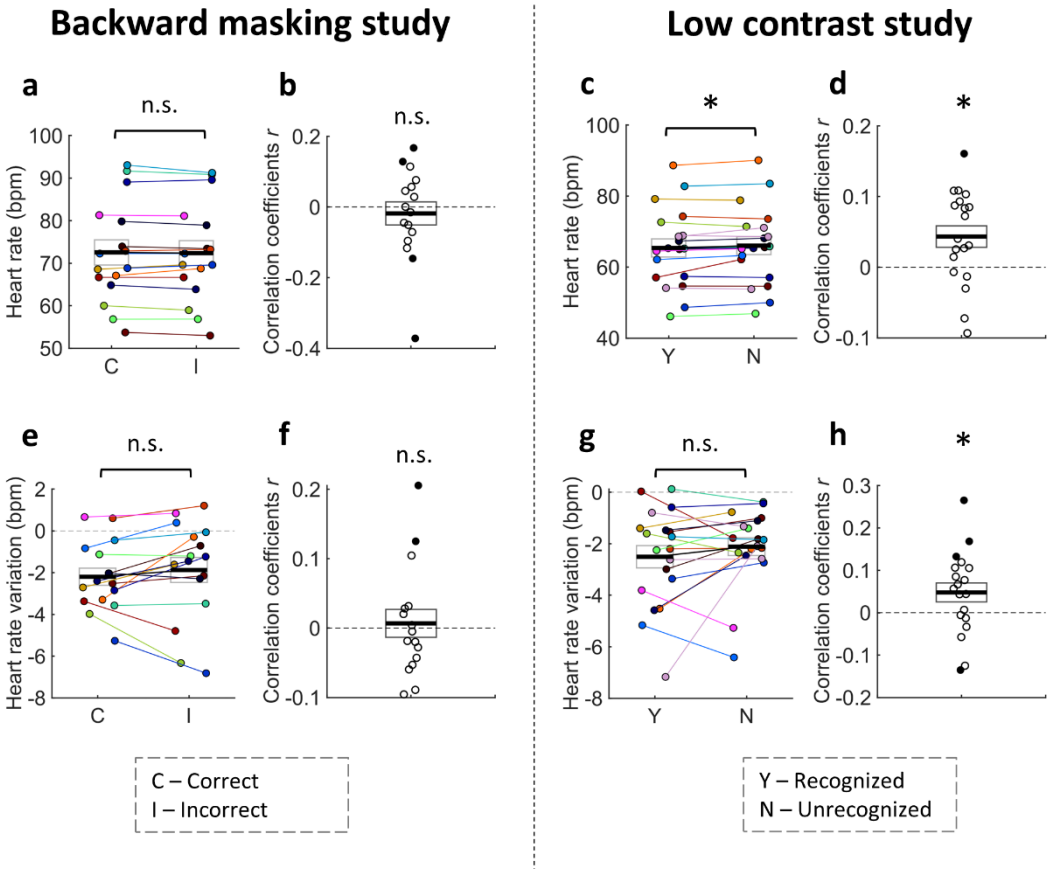


Figure 3.4 *– Fluctuations in heart rate, measured in the cardiac cycle just before stimulus onset, are associated with visual performance.* ***a*** *Relation between heart rate and participant's accuracy (backward masking study).* ***b*** *Correlation between heart rate and RT (backward masking study).* ***c*** *Relation between heart rate and participant's recognition (low contrast study).* ***d*** *Correlation between heart rate and RT (low contrast study).* ***e*** *Relation between heart rate variation and participant's accuracy (backward masking study).* ***f*** *Correlation between heart rate variation and RT (backward masking study).* ***g*** *Relation between heart rate variation and participant's recognition (low contrast study).* ***h*** *Correlation between heart rate variation and RT (low contrast study). In sub-figures* ***a, c, e*** *and* ***g*** *the black horizontal line is the average heart rate/heart rate variation and the rectangle represents ± standard error of the mean. In sub-figures* ***b, d, f*** *and* ***h*** *the black line is the average correlation coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. Filled circles represent participants where correlation is statistically significant (p<0.05). *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.1.2.2 Phase of the cardiac cycle at the time of visual stimulus onset

Lastly, we repeated all the analyses described above but with the phase of cardiac cycle at the time of visual stimulus onset. This decision was motivated by previous studies whose results suggest that the perception of a tactile stimulus is facilitated when it is presented during diastole compared to when it is presented in systole (32). So, we hypothesized a better performance or an easier detection of the stimulus in the trials in which the stimulus appeared in the diastolic phase.

The cardiac cycle has a duration of approximately 700ms, with 270ms corresponding to systole and the remaining 430ms corresponding to diastole (126). Considering the duration of each phase of the cardiac cycle, we can convert the duration of each phase to degrees and thus analyze the circular component of the cycle. Applying a simple rule, systole can be estimated to be around 140º, while diastole is approximately 220º. It is to be noted that we determined as the beginning of the cycle the S peak and that the systole begins before, closer to the R peak. That said, we can consider that systole, in terms of the trigonometric circle, corresponds to the end of the fourth quadrant, the total of the first quadrant and the initial part of the second quadrant. Since we used the sine and cosine of the phase, considering the circular component of the cycle, we hypothesized that, when the stimulus occurred in the mentioned quadrants, the detection would be lower. Consequently, looking at the trigonometric circle, we anticipated that, in incorrect trials, as well as in unrecognized trials, the stimulus would be presented mostly in the first quadrant. The stimulus could also appear at the end of the fourth quadrant and at the beginning of the second quadrant, which would result in correspondingly negative sine and cosine values. However, these negative values would be very close to zero. Having said that, it was expected that the average of sine and cosine values would be positive. To illustrate this concept, we expected that, in both incorrect trials of the backward masking study and unrecognized trials of the low contrast study, the stimulus would predominantly appear in the region delineated by the two orange lines (Figure 3.5 – a)).

As previously conducted, let's start by the backward masking study. Looking to Figure 3.5 - b), we observe a more positive value for the sine component for incorrect trials, which goes according to what we were expecting. However, to the cosine component we see an average mean value more negative (Figure 3.5 - c)). Thus, on average, in incorrect trials, the stimulus appears outside the previously delimited region. Anyway, the difference between correct and incorrect trials is not statistically significant (sin θ - Shapiro-wilk test: $W = 0.973$, $p = 0.889$ and paired-sample t-test: $t(15) = -1.257$, $p = 0.228$; cos θ - Shapiro-wilk test: $W = 0.932$, $p = 0.258$ and paired-sample t-test: $t(15) = 1.2316$, $p = 0.237$). These results contradict those found in (32), which suggest that performance is higher when the stimulus is presented at a given moment of the cardiac cycle, namely diastole.

In the low contrast study, there are no significant differences between trial types regarding cosine component (Figure 3.5 – f); Shapiro-wilk test: $W = 0.925$, $p = 0.123$ and paired-sample t-test: $t(18) = -0.981$, $p = 0.340$). In the sine component, in turn, we can see significant differences between recognized and unrecognized trials (Figure 3.5 – e); Shapiro-wilk test: $W = 0.940$, $p = 0.225$ and paired-sample t-test: $t(18) = 2.506$, $p = 0.022$). As mentioned, we expected that, in unrecognized trials, the stimulus would predominantly appear in the region delineated by the two orange lines, which corresponds to systole. Thus, it would be anticipated that the average of the sine component would be positive. However, this is not

what happens. We see that, in the unrecognized trials, the sine component is predominantly negative. This result suggests that the recognition of the stimulus is lower when the stimulus is presented in the third and fourth quadrant of the trigonometric circle, which, in the cardiac cycle, corresponds to diastole. These findings contradict those found in (32), which suggest that performance is higher when the stimulus is presented at a given moment of the cardiac cycle, namely diastole. Therefore, although a relatively small relationship, in the low contrast study, cardiac phase appears to affect performance.

Similar to the analysis conducted on other cardiac measures, we also correlate the phase of the cycle at the time of stimulus presentation with the RT. To do that, we did a regression, which allowed us to obtain the regression coefficients for both the cosine and the sine of the phase. Once again, the analysis was performed for both studies, but none of them found regression coefficients significantly different from zero (Figure 3.5 - d) and Figure 3.5 - g)). In the backward masking study the results were: sin θ - Shapiro-wilk test: W = 0.907, p = 0.095 and one-sample t-test: $t(15)$ = -0.754, p = 0.462; cos θ - Shapiro-wilk test: W = 0.879, p = 0.038 and Wilcoxon test: Z = 0.983, p = 0.326. For the low contrast study, the results were: sin θ - Shapiro-wilk test: W = 0.826, p = 0.004 and Wilcoxon test: Z = 1.207, p = 0.227; cos θ - Shapiro-wilk test: W = 0.010, p = 0.011 and Wilcoxon test: Z = -1.123, p = 0.260. The results obtained suggest that the cardiac phase does not influence the RT, i.e., there seems to be no significant relationship between the cardiac phase and the time required to make the decision.

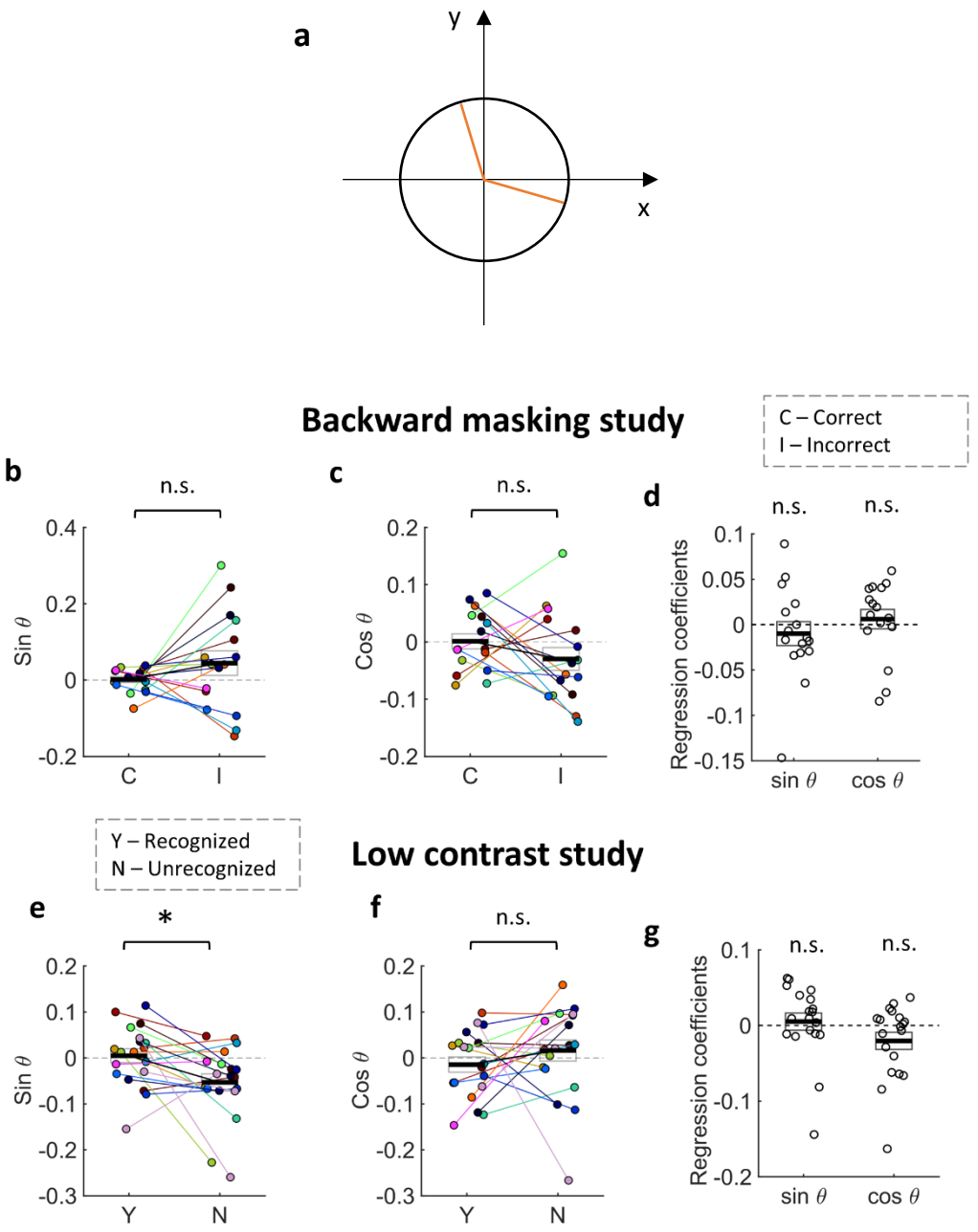


Figure 3.5 *- The phase of the cardiac cycle in which the stimulus appears is associated with visual performance in the low contrast study.* ***a*** *Possible representation of systole in the trigonometric cycle - systole corresponds to the circle zone between the orange lines.* ***b*** *Relation between the sine component of the phase and participant's accuracy (backward masking study).* ***c*** *Relation between the cosine component of the phase and participant's accuracy (backward masking study).* ***d*** *Regression between sine and RT and between cosine and RT, respectively (backward masking study).* ***e*** *Relation between the sine component of the phase and participant's recognition (low contrast study).* ***f*** *Relation between the cosine component of the phase and participant's recognition (low contrast study).* ***g*** *Regression between sine and RT and between cosine and RT, respectively (low contrast study). In sub-figures* ***b, c, e*** *and* ***f*** *the black horizontal line is the average sine/cosine and the rectangle represents ± standard error of the mean. In sub-figures* ***d*** *and* ***g*** *the black horizontal line is the average regression coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.1.3 SVM: classification of visual performance based on pre-stimulus cardiac activity

Upon analyzing Figure 3.6 - a), we can observe that only the heart rate measure allows us to predict whether the participants correctly identified the stimulus category in the backward masking study. This conclusion is supported by the results in Table 8, where we can see that, when using heart rate measure, 55% of the classifiers surpass random performance. Conversely, in the remaining measures, the percentage of classifiers with a superior performance to the random behavior is lower than 10%. Continuing the analysis of Figure 3.6 - a), we can see that the combination of all measures does not allow predicting the accuracy. We expected a different behavior, since with the heart rate the classifier can extract activity patterns. By combining all the measures, even though the phase and heart rate variation do not contain enough information, we expected that the classifier would be able to extract the heart rate patterns that it extracted when studying this measure alone. However, this was not the observed behavior. We believe that this observed behavior can be explained by conflicting information within the different measures, which makes it difficult for the classifier to perceive which patterns are associated with correct and incorrect trials. Consequently, we can deduce that among all the measures, the heart rate is the sole one that provides pertinent information to the classifier.

Subsequently, the same analysis was performed for the low contrast study. Table 8 reveals that only some of the classifiers perform better than the random behavior, with a maximum of 25% observed in the study of each measure separately. That said, we can deduce that, within this study, cardiac activity does not provide enough information for the classifier to predict whether the participant recognized the stimulus. Consequently, we proceeded to study the influence of each measure individually. In Figure 3.6 - b), we see no significant differences between the classifier's performance when using each measure separately and when combining all measures. Consequently, we can conclude that, in the low contrast study, pre-stimulus cardiac activity does not facilitate the prediction of participant's recognition.

The results obtained in both studies are contradictory. As observed earlier, the low contrast study demonstrates a notable and significant difference in heart rate between recognized and unrecognized trials, whereas such difference is missing when comparing correct and incorrect trials in the backward masking study. That said, it would be expected that in the low contrast study it would be possible to predict recognition through cardiac activity, whereas in the backward masking study it would not be possible. However, the opposite behavior is verified. The results suggest that the classifier was able to extract activity patterns in heart rate in the backward masking study that were not detectable through traditional statistical analysis.

	Backward masking study	Low contrast study
All cardiac measures	0%	15%
Heart rate	**55%**	25%
Heart rate variation	0%	25%
Phase	5%	25%

Table 8 *– Percentage of classifiers using cardiac measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	Low contrast study
Heart rate	***t (19)* = -10.405; *p* < 0.001**	*t (19)* = -1.564, *p* = 0.134
Heart rate variation	***t (19)* = 2.390; *p* = 0.027**	*t (19)* = -1.794, *p* = 0.089
Phase	*t (19)* = -1.886; *p* = 0.075	*t (19)* = -1.726, *p* = 0.101

Table 9 *– Comparison between the use of each cardiac measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all cardiac measures were combined. The statistically significant analyses are highlighted.*

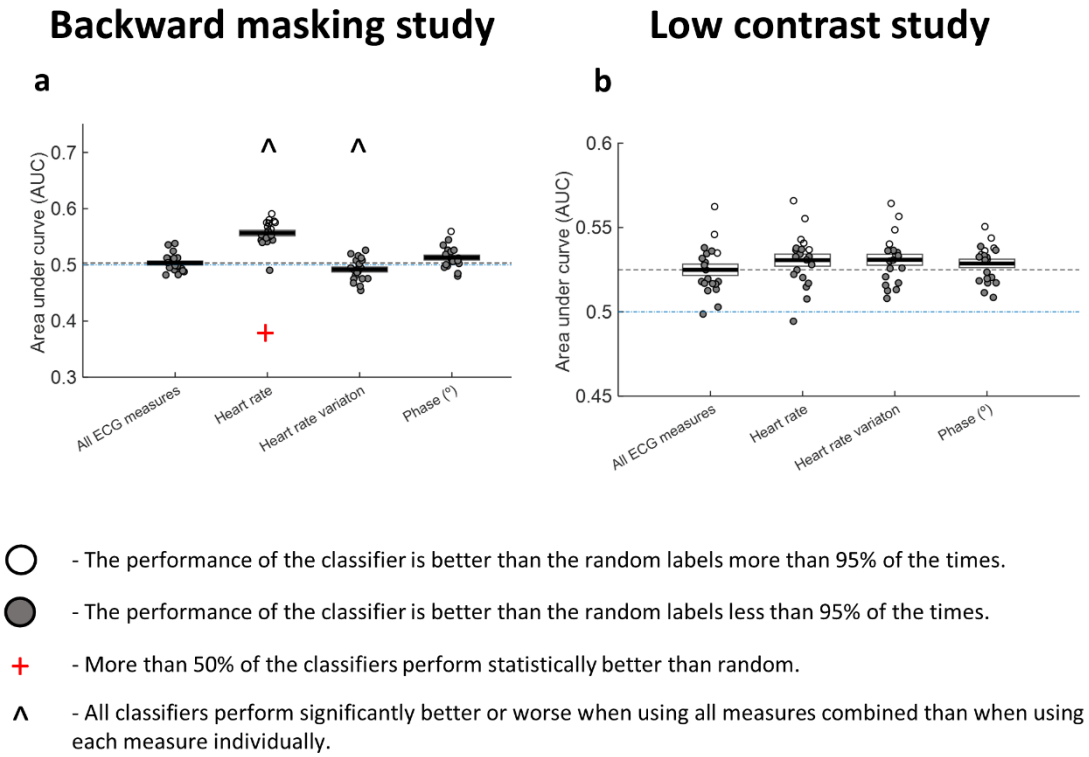


Figure 3.6 *– Pre-stimulus cardiac activity allows predicting participant's accuracy in the backward masking study but does not allow predicting stimulus recognition in the low contrast study.* ***a*** *AUC using a SVM model in the backward masking study.* ***b*** *AUC using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all cardiac measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed. +: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random. ^: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

3.2.2 Effect of pre-stimulus respiratory activity on visual perception

3.2.2.1 Study of the respiratory cycle duration variation evoked by the warning cue

We found that, in both studies, the auditory cue is followed by an increase in the duration of the respiratory cycle and then followed by a decrease (Figure 3.7 – a) and Figure 3.7 - c)). Permutation tests revealed that the cycle duration that includes the visual stimulus was significantly increased in comparison with the duration of the cycle before the cue. In Figure 3.7 – b) and Figure 3.7 - d), we can see that the maximum of cycle duration modulation occurred in the cycle that contains the visual stimulus. This observation suggests that

sensorimotor processing is associated with an increase in respiratory cycle duration. Following that, our results suggest an acceleration in respiratory rhythm, i.e., decrease in cycle duration. This finding suggests that the presentation of visual stimulus induces a deceleration in breathing activity, maybe as an attempt to facilitate visual processing or maybe corresponds to the return of the normal state, which had been interrupted by the induction of the state of expectation. Comparing the two studies, we can see that, in the low contrast study, cycle length restoration takes longer than in the backward masking study. This difference might be explained by task variations that, consequently, influence visual processing.

Along these lines, we also conducted an analysis of the variation in cycle duration between correct/incorrect trials (backward masking study) and recognized/unrecognized trials (low contrast study). In the backward masking study, we can see significant differences only in the fourth cycle after cue onset (Figure 3.8 - a)), which appears after visual stimulus onset. This result suggests that the observed difference is not cue induced. Finally, in the low contrast study, there are no significant differences between the trial types (Figure 3.8 - c) and Figure 3.8 – d)).

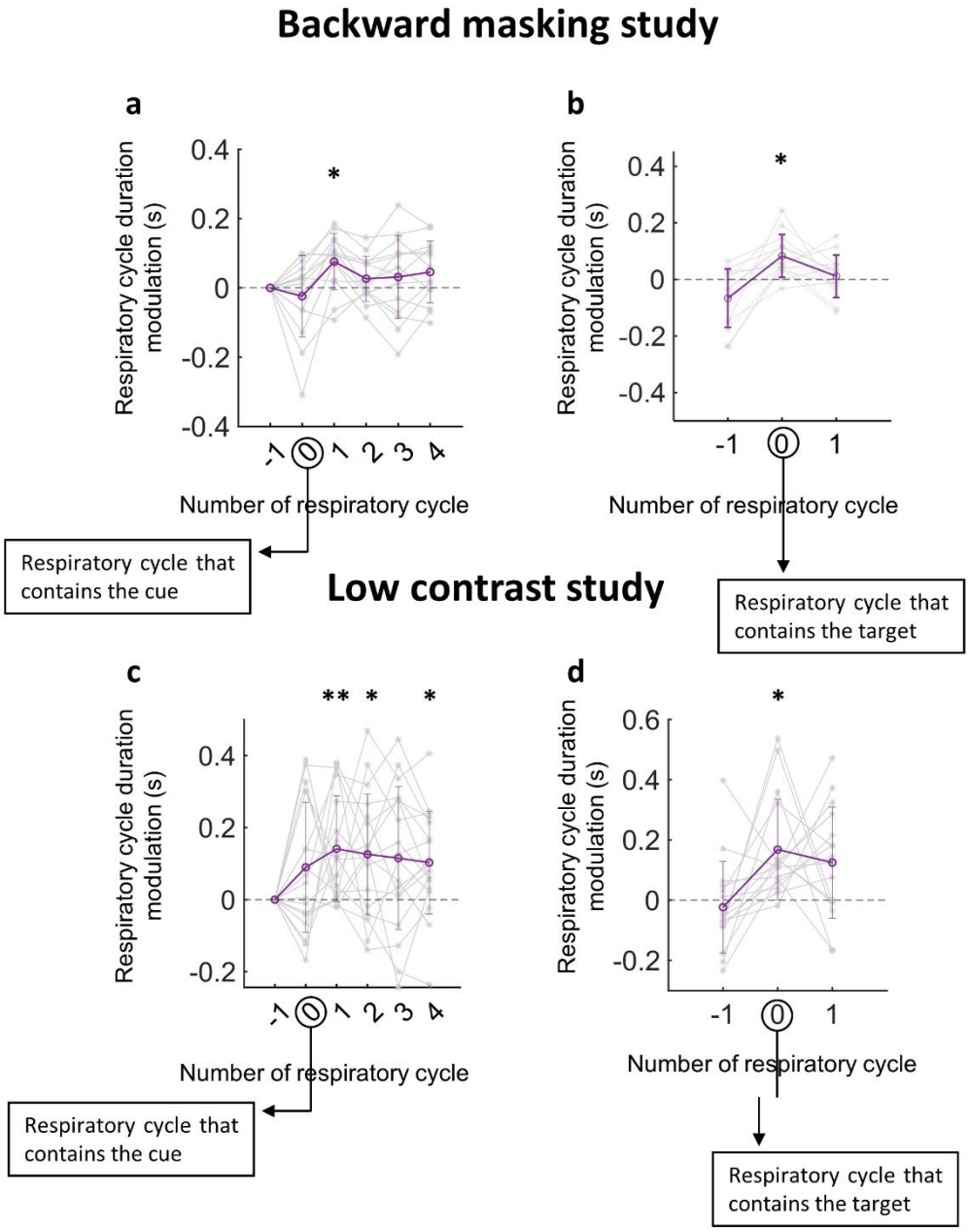


Figure 3.7 *– Sensorimotor processing is associated with an increase in the duration of the respiratory cycle.* ***a*** *Variation of respiratory cycle duration in one cycle before and four cycles after cue onset (backward masking study).* ***b*** *Variation of respiratory cycle duration in the cycle before and after target onset (backward masking study).* ***c*** *Variation of respiratory cycle duration in one cycle before and four cycles after cue onset (low contrast study).* ***d*** *Variation of respiratory cycle duration in the cycle before and after target onset (low contrast study). In sub-figures* ***a*** *and* ***c****, the zero on the x-axis represents the respiratory cycle that contains the auditory cue. In sub-figures* ***b*** *and* ***d****, the zero on the x-axis represents the respiratory cycle that contains the target. In all graphs, data are represented as mean ± standard error of the mean across participants. The asterisks signal the cycles where cycle duration variation is significantly different from zero (*p<0.05, **p≤0.01, ***p≤0.001)*

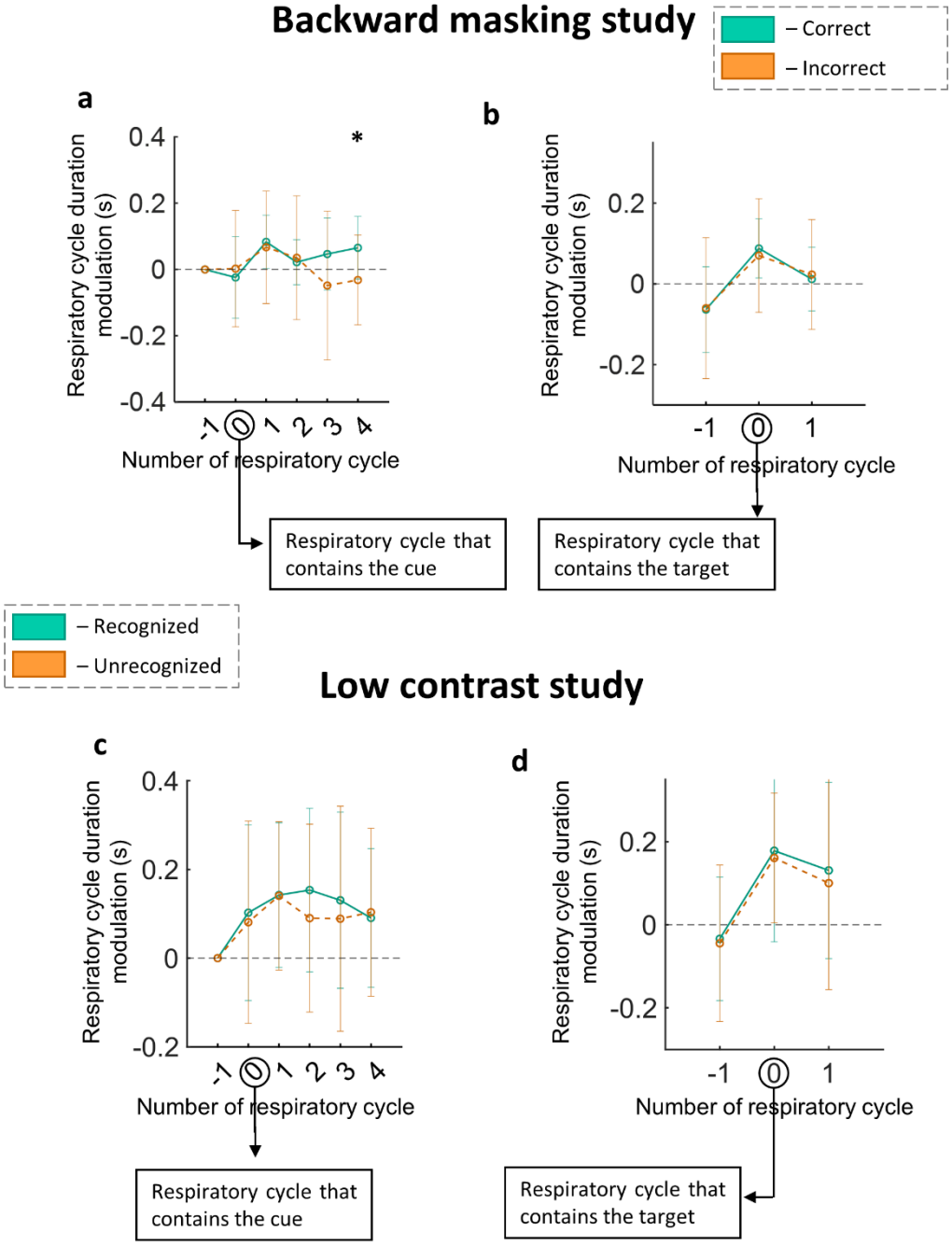


Figure 3.8 *- There are no significant differences in pre-stimulus respiratory activity between correct and incorrect trials (backward masking study) and recognized and unrecognized trails (low contrast study).* ***a*** *Variation in the duration of the respiratory cycle in correct and incorrect trials in one cycle before and four cycles after cue onset (backward masking study).* ***b*** *Variation in the duration of the respiratory cycle in correct and incorrect trials in the cycle before and after target onset (backward masking study).* ***c*** *Variation in the duration of the respiratory cycle the recognized and unrecognized trials in one cycle before and four cycles after cue onset (low contrast study).* ***d*** *Variation in the duration of the respiratory cycle in correct and incorrect trials in the cycle before and after target onset (low contrast study). In sub-figures* ***a*** *and* ***c****, the zero on the x-axis represents the respiratory cycle that contains the auditory cue. In sub-figures* ***b*** *and* ***d****, the zero on the x-axis represents the respiratory cycle that contains the target. In all graphs, data are represented as mean ± standard error of the mean across participants. The asterisks signal the cycles where cycle duration variation is significantly different between trial types (*p<0.05, **p≤0.01, ***p≤0.001).*

3.2.2.2 Analyses of respiratory features used in the classifiers

For the analyses of the effect of pre-stimulus respiratory activity in visual perception using classifiers, we extracted two different types of features: respiratory cycle duration and respiratory phase. It should be noted that, to obtain the respiratory cycle duration, we extracted the length of the cycle just before stimulus onset. First, we studied how these were related to task performance in both studies and then we investigated if an SVM classifier was able to use them to predict the participant's performance.

3.2.2.2.1 Respiratory cycle duration

In both Figure 3.9 – a) and Figure 3.9 – c), we can see that the comparison of the respiratory cycle duration between the trial types in both studies (correct/incorrect trials in the backward masking study and recognized/unrecognized trials in the low contrast study) revealed no significant differences (backward masking study: Shapiro-wilk test: $W = 0.913$, $p = 0.150$ and paired-sample t-test: $t(13) = 0.284$, $p = 0.781$; low contrast study: Shapiro-wilk test: $W = 0.698$, $p < 0.001$ and Wilcoxon test: $Z = 1.023$, $p = 0.306$). Considering that, in both studies, there is no relation between visual stimulus performance and cycle duration, we can hypothesize that the respiratory rate is not associated with visual performance and, probably, does not influence visual processing.

Similar to the analyses conducted for other measures, we also explored the correlation between respiratory cycle duration and RT. Since, in the previous analysis, none of the studies showed significant differences between the trial types, we expected that a correlation between RT and respiratory cycle would not be observed. In fact, the results go according to what we were expecting, since in none of the studies the correlation values are statistically different from zero (backward masking study: Shapiro-wilk test: $W = 0.955$, $p = 0.639$ and one-sample t-test: $t(13) = 0.550$, $p = 0.591$; low contrast study: Shapiro-wilk: $W = 0.934$, $p = 0.230$ and one-sample t-test: $t(17) = -1.075$, $p = 0.297$) (Figure 3.9 – b) and Figure 3.9 – d)) . The findings obtained from this analysis are in alignment with previous observations. In none of the analyses there seems to be an association between the respiratory cycle duration and the visual processing.

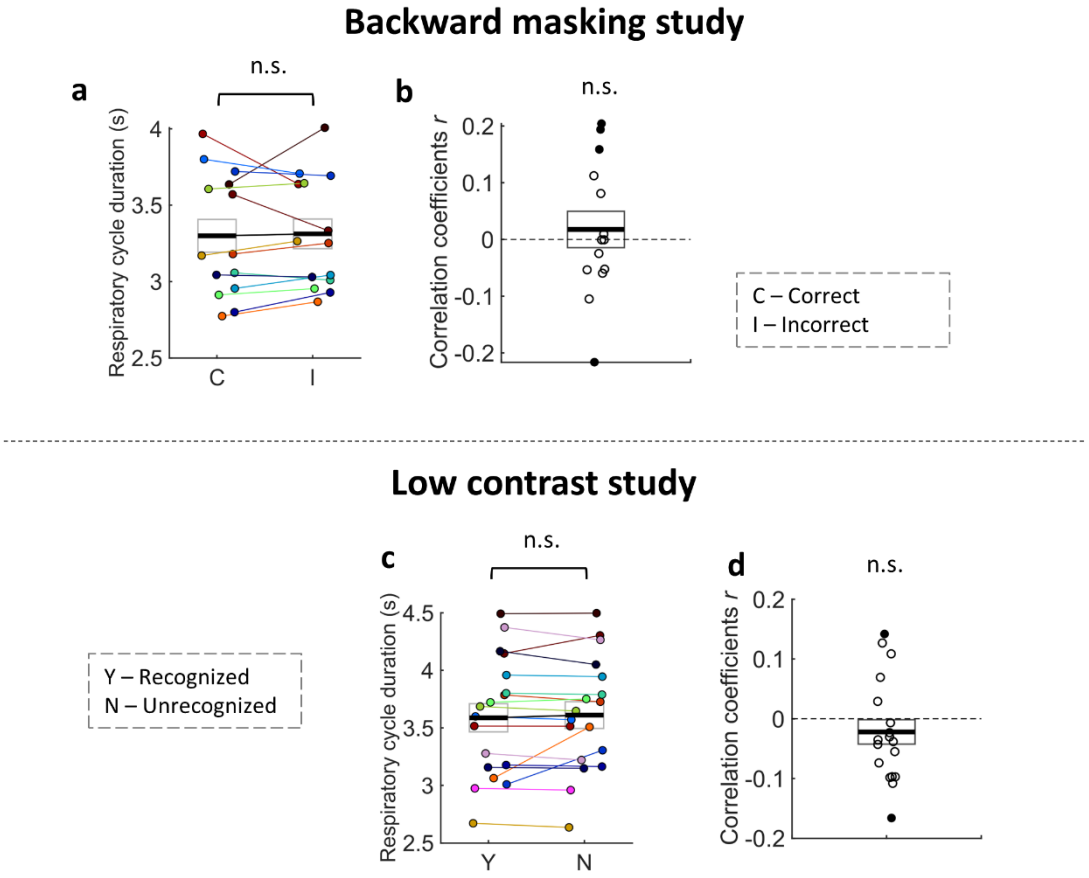


Figure 3.9 – *The respiratory rhythm seems not to be associated with visual performance. The respiratory cycle duration was measured considering the respiratory cycle just before stimulus onset.* ***a*** *Relation between respiratory cycle duration and participant's accuracy (backward masking study).* ***b*** *Correlation between respiratory cycle duration and RT (backward masking study).* ***c*** *Relation between respiratory cycle duration and participant's recognition (low contrast study).* ***d*** *Correlation between respiratory cycle duration and RT (low contrast study). In sub-figures* ***a*** *and* ***c*** *the black horizontal line is the average respiratory cycle duration and the rectangle represents ± standard error of the mean. In sub-figures* ***b*** *and* ***d*** *the black horizontal line is the average correlation coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. Filled circles represent participants where correlation is statistically significant (p<0.05). *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant*

3.2.2.2.2 Phase of respiratory cycle at the time of visual stimulus onset

In 2019, Perl, O. et al. saw that when the stimulus is presented in synchronization with inhalation, there is an optimization of stimulus processing, which results in a higher performance (4). Hence, we expected to see a better performance or an easier detection in the trials in which the stimulus appeared in inhalation. Considering, as explained in the methods section, that we defined the beginning of the cycle as the moment in which a valley occurs, we can assume that the beginning of the cycle corresponds to the beginning of inhalation. In this way, we hypothesized that stimulus detection, as well as accuracy, would be higher when the stimulus is presented in the beginning of the cycle. The ratio between the inspiratory and

expiratory phase is 1:2, indicating that the duration of the expiratory phase is twice as long as the inspiratory phase (127). By converting this proportion to degrees, we can establish that the inspiratory phase corresponds to approximately 120º, while the expiratory phase accounts for the remaining 240º. That said, we can consider that inspiratory phase, in terms of the trigonometric circle, corresponds to the first quadrant and to the initial portion of the second quadrant. Since we used the sine and cosine of the phase, considering the circular component of the cycle, we hypothesized that, when the stimulus occurred in the mentioned quadrants, the detection would be higher. Consequently, looking at the trigonometric circle, we anticipated that, in correct trials, as well as in recognized trials, the stimulus would be presented mostly in the first quadrant. The stimulus could also appear at the beginning of the second quadrant, which would result in correspondingly negative cosine values. However, these negative values would be very close to zero. Having said that, it was expected that the average of sine and cosine values would be positive. To illustrate this concept, we expected that, in both correct trials of the backward masking study and recognized trials of the low contrast study, the stimulus would predominantly appear in the region delineated by the two orange lines (Figure 3.10 – a)).

As previously conducted, let us begin with the backward masking study. Looking to Figure 3.10 - b) and Figure 3.10 – c), we observe more positive values in both sine and cosine components for incorrect trials, which contradicts our initial expectation. Positive sine and cosine values were expected for correct trials and not for incorrect ones. Applying statistical tests, we found no significant difference between correct and incorrect trials (sin θ - Shapiro-wilk test: $W = 0.933$, $p = 0.333$ and paired-sample t-test: $t(13) = -0.652$, $p = 0.526$; cos θ - Shapiro-wilk test: $W = 0.947$, $p = 0.458$ and paired-sample t-test: $t(13) = -0.094$, $p = 0.927$). Based on the results of these statistical tests, we can conclude that the phase of the respiratory cycle did not significantly influence the participant's performance in the backward masking study.

The identical analytical approach was employed in the low contrast study. Again, we anticipated observing more positive sine and cosine values in the recognized trials. Looking at Figure 3.10 – e), we can see a higher, but not significant, value of sine component for recognized trials compared to unrecognized trials; however, the average value is negative (Shapiro-wilk test: $W = 972$, $p = 0.832$ and paired-sample t-test: $t(17) = 0.467$, $p = 0.646$). In Figure 3.10 – f), we also see no significant patterns in the cosine component, being the average value in recognized trials practically zero (Shapiro-wilk test: $W = 0.956$, $p = 0.590$ and paired-sample t-test: $t(17) = 0.790$, $p = 0.440$). In this way, the sine and cosine components in recognized trials are not predominantly positive, which suggest that the inspiratory phase is not the phase that favors stimulus detection. Considering the results, we can conclude that, in the low contrast study, the phase of the respiratory cycle is not associated with participant's recognition.

Similarly to the analysis conducted on other respiratory measures, we also examined the phase of the cycle at the time of stimulus presentation with the RT. To achieve this, we performed a regression, which allowed us to obtain the regression coefficients for both the cosine and the sine of the phase. Once again, these analyses were performed for both studies, but none of them found regression coefficients significantly different from zero (Figure 3.10 – d) and Figure 3.10 – g)). In the backward masking study, the results were: sin θ - Shapiro-wilk test: $W = 0.950$, $p = 0.561$ and one-sample t-test: $t(13) = -0.032$, $p = 0.974$; cos θ - Shapiro-wilk test: $W = 0.862$, $p = 0.083$ and one-sample t-test: $t(13) = 0.802$, $p = 0.467$. For the low contrast study, the results were: sin θ - Shapiro-wilk test: $W = 0.978$, $p = 0.928$ and one-sample t-test: $t(17) = -0.867$, $p = 0.398$; cos θ - Shapiro-wilk test: $W = 0.958$, $p = 0.554$ and one-sample t-test: $t(17) = 0.1367$, $p = 0.893$.

Taking into consideration all the findings in this section, we speculate that the phase of the respiratory cycle in which the stimulus is presented is not associated with visual stimulus recognition.

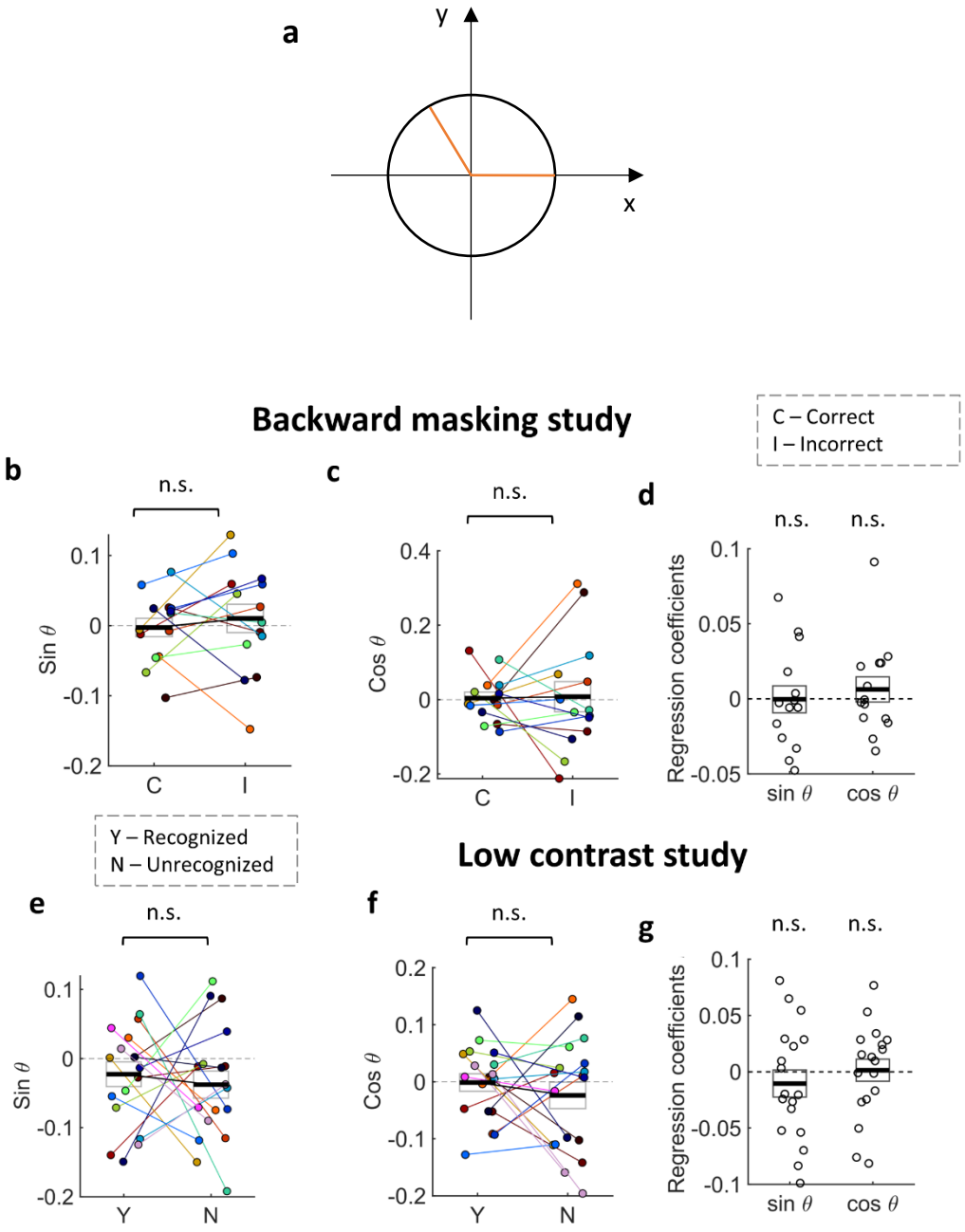


Figure 3.10 *– The phase of the respiratory cycle in which the stimulus appears does not influence its perception.* ***a*** *Possible representation of inspiration in the trigonometric cycle – inspiration corresponds to the circle zone between the orange lines.* ***b*** *Relation between the sine component of the phase and participant's accuracy (backward masking study).* ***c*** *Relation between the cosine component of the phase and participant's accuracy (backward masking study).* ***d*** *Regression between sine and RT and between cosine and RT, respectively (backward masking study).* ***e*** *Relation between the sine component of the phase and participant's recognition (low contrast study).* ***f*** *Relation between the cosine component of the phase and participant's recognition (low contrast study).* ***g*** *Regression between sine and RT and between cosine and RT, respectively (low contrast study). In sub-figures* ***b***, ***c***, ***e*** *and* ***f*** *the black horizontal line is the average sine/cosine and the rectangle represents ± standard error of the mean. In sub-figures* ***d*** *and* ***g*** *the black horizontal line is the average regression coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.2.3 SVM: classification of visual performance based on pre-stimulus respiratory activity

In a similar manner to cardiac measures, we also investigated whether there are patterns in respiratory activity that may facilitate stimulus detection and task performance. In Table 10, we can see that, in the backward masking study, when we applied the permutation tests, only the parameter related to respiratory cycle duration exhibits performance significantly superior to random behavior. In other words, this particular parameter is the only one with at least 50% of classifiers demonstrating better performance than random behavior. Conversely, the respiratory phase parameter does not appear to have enough information to allow the classifier to predict trial performance, since only 40% of the classifiers demonstrate better performance than random. In Figure 3.11 - a), we can see that by combining the features of respiratory activity there is a deterioration in classifier performance. This observation suggests, in some way, that these measures might contain conflicting information. In conclusion, this result may come from the fact that the backward masking study is very noisy.

In the low contrast study, the classifiers were able to extract activity patterns from each combination of features allowing predicting stimulus recognition. From Table 10, it becomes evident that in all combinations of features employed, over 70% of the classifiers achieved a performance superior to the random behavior. We then studied the influence of each parameter individually. In Figure 3.11 – b), we can see that the performance of each feature separately is almost the same as the performance of all features combined. This result suggests that the combination of all measures does not provide extra information to the classifier.

As no significant differences were found in the previous analyses, we did not expect the classifier to be able to predict visual performance. In the backward masking study, only the respiratory cycle duration parameter offers relevant information to the classifier. However, even within this context, only 55% of the classifiers were able to outperform random, which is very close to the threshold we established. In contrast, in the low contrast study, it was possible to predict stimulus recognition through respiratory activity. In this study, classifiers were able to recognize patterns of activity that were not detected in previous analyses. We believe that this disparity may be attributed to the fact that our previous analyses focused on the average behavior of each participant and not on the activity pattern across trials.

	Backward masking study	Low contrast study
All respiratory measures	10%	**80%**
Respiratory cycle duration	**55%**	**70%**
Phase	40%	**70%**

Table 10 – *Percentage of classifiers using respiratory measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	Low contrast study
Respiratory cycle duration	***t (19)* = -2.898; *p* = 0.009**	*t (19)* = -1.267; *p* = 0.220
Phase	*t (19)* = -1.825; *p* = 0.084	*t (19)* = -0.467; *p* = 0.646

Table 11 *– Comparison between the use of each respiratory measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all respiratory measures were combined. The statistically significant analyses are highlighted.*

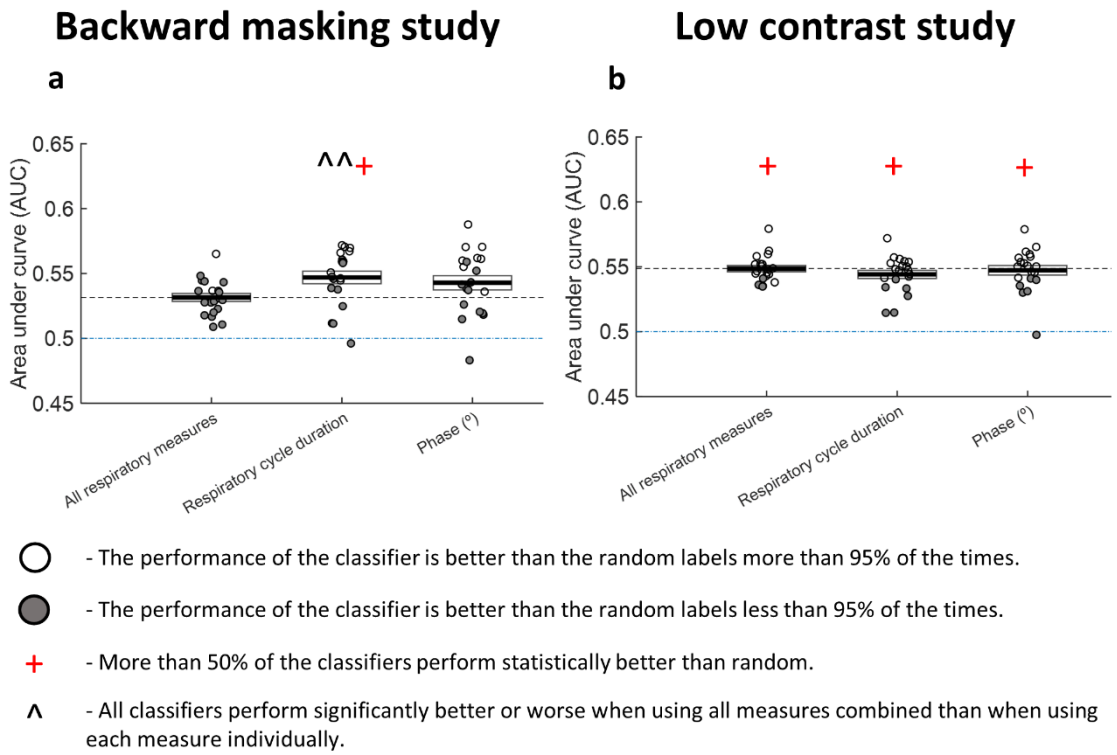


Figure 3.11 *- Pre-stimulus respiratory activity allows predicting the participant's recognition.* ***a*** *AUC using a SVM model in the backward masking study.* ***b*** *AUC using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all respiratory measures is used. The remaining bars*

concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed. +: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random. ^: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.

3.2.3 Effect of pre-stimulus pupillary response

3.2.3.1 Study of the time course of the pupillary response evoked by the warning cue

Recent research findings suggest that the state of expectation among other things activates the sympathetic system, which results in pupil dilation (82). Allied to this, prior research demonstrates that greater pupil dilation suggests increased cognitive activity, implying that in such instances stimulus detection tends to be more favorable (128). In Figure 3.12 - a) and Figure 3.12 - b), these conclusions are supported. We can see that, initially, the auditory signal induces a constriction followed by significant dilation. Contrary to what is seen in the backward masking study, in the low contrast study, the constriction is more pronounced and has a longer duration, i.e., the pupil starts to dilate later (about three seconds after cue onset). Comparing the two studies, we note that, in the backward masking study, the maximum magnitude of pupil response is higher. We hypothesize that the changes observed in pupil behavior may be attributed to variations in task difficulty. As the two tasks differ from each other, it is plausible to assume that the required level and type of concentration needed may also differ. In this way, the participants prepare themselves differently for the two tasks, which may result in differences in pupil behavior.

In Figure 3.12 – c), Figure 3.12 – d), Figure 3.12 – g) and Figure 3.12 – h), we can see that after stimulus onset there is a substantial increase in pupil size, resulting in a distinct pupil dilation. This dilation phenomenon started approximately between 300 to 500ms, which corresponds to the decision-making processes. After that, the dilation extended for approximately one second after stimulus onset, coinciding with the point at which the subject made a choice and provided a response with the chosen category option (in the task performance section, we can see that the average reaction time in both studies is approximately one second after stimulus onset - Figure 3.2).

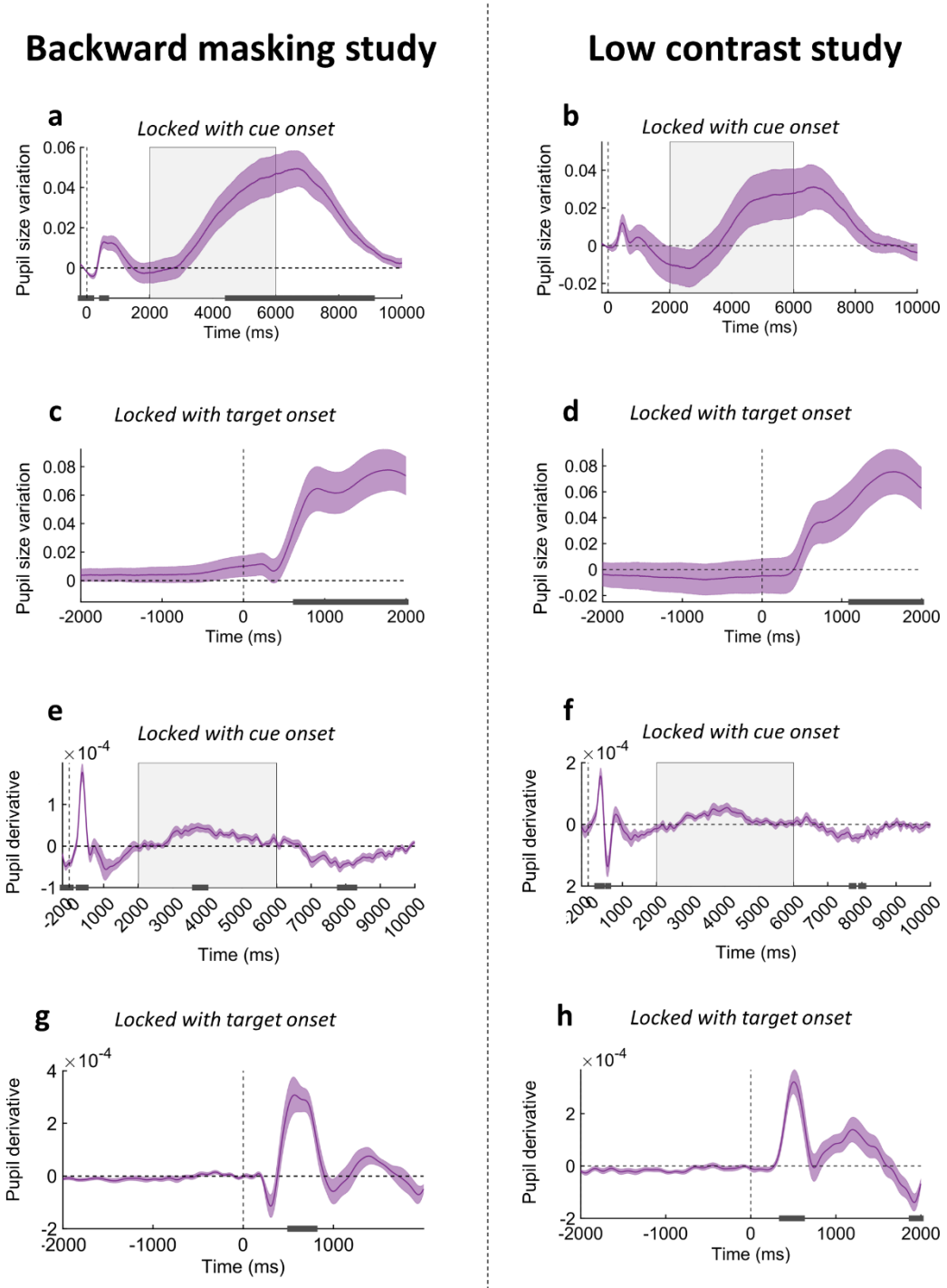


Figure 3.12 *– The cue induced pupil dilation.* ***a and b*** *Cue-locked pupil size modulation (a - backward masking study; b - low contrast study).* ***c and d*** *Target-locked pupil size modulation (a - backward masking study; b - low contrast study).* ***e and f*** *Cue-locked pupil derivative modulation (a - backward masking study; b - low contrast study).* ***g and h*** *Target-locked pupil derivative modulation (a - backward masking study; b - low contrast study). The gray horizontal line represents the significant time points where the pupil size modulation or the pupil derivative modulation is significantly different from zero ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. In all graphs, data are represented as mean ± standard error of the mean across participants.*

In the backward masking study, when comparing correct with incorrect trials, we see no significant differences in pupillary response (Figure 3.13 – a), Figure 3.13 – c), Figure 3.13 – e) and Figure 3.13 – g)). It was expected that, in the correct trials, the dilation would be more pronounced, which would produce a more pronounced and positive behavior in the derivative. For the low contrast study, looking at Figure 3.13 – b) and Figure 3.13 – f), it is apparent that there is no significant difference in pupillary response between recognized and unrecognized trials. However, Figure 3.13 – d) and Figure 3.13 – h) reveal distinct times points where the difference between the two conditions is statistically significant. It should be noted that these significant time points occur after stimulus onset, so the differences were not induced by the state of expectation. These statistical differences may be linked to the decision-making process.

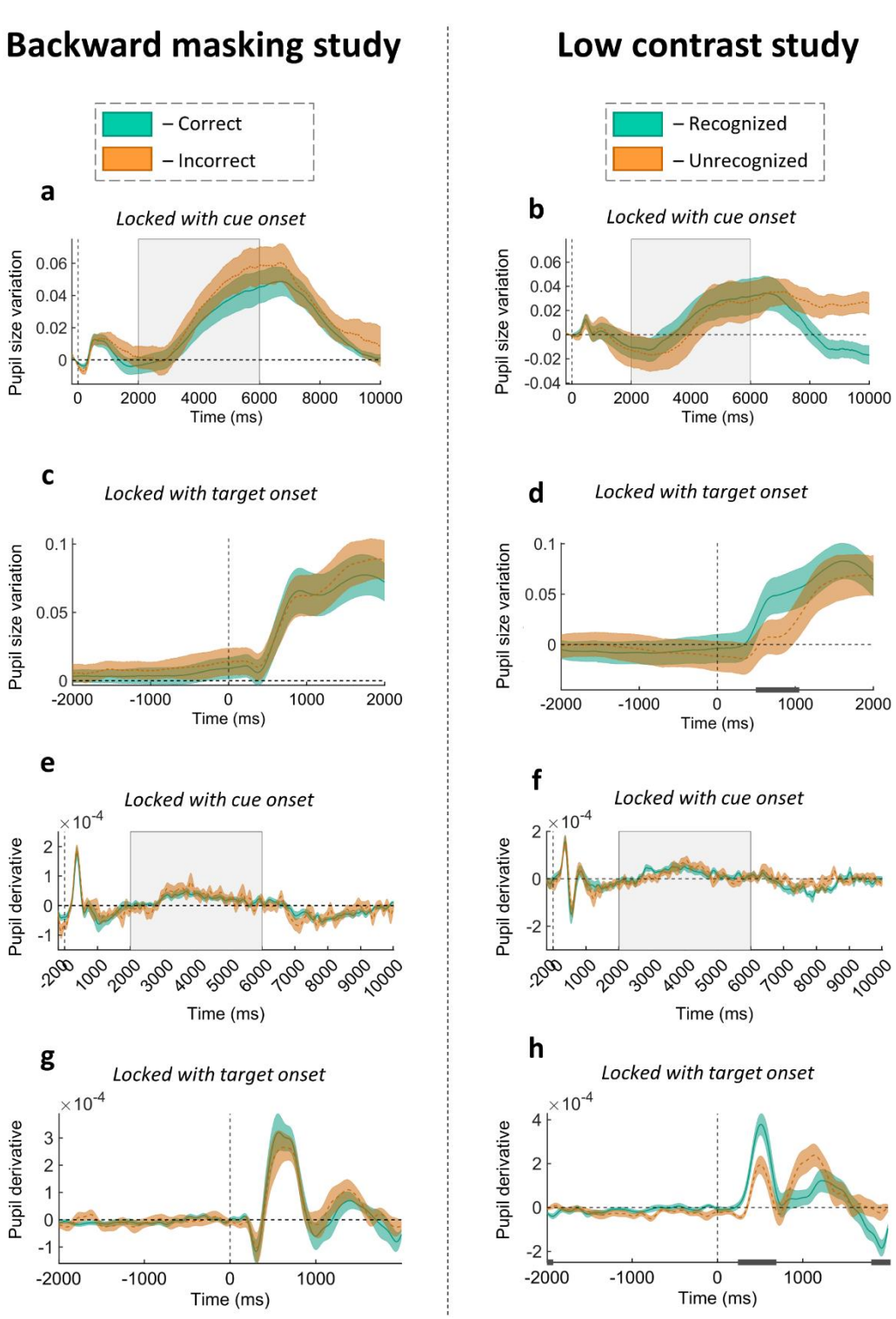


Figure 3.13 *– The state of expectation did not induce significant differences in pupillary response between correct and incorrect trials (backward masking study) or between recognized and unrecognized trials (low contrast study).* ***a*** *Cue-locked pupil size modulation in correct (green curve) and in incorrect trials (orange curve) (backward masking study).* ***b*** *Cue-locked pupil size modulation in recognized (green curve) and unrecognized trials (orange curve) (low contrast study).* ***c*** *Target-locked pupil size modulation in correct (green curve) and in incorrect trials (orange curve) (backward masking study).* ***d*** *Target-locked pupil size modulation in recognized (green curve) and in unrecognized trials (orange curve) (low contrast study).* ***e*** *Cue-locked pupil derivative modulation in correct (green curve) and in incorrect trials (orange curve) (backward masking study).* ***f*** *Cue-locked pupil derivative modulation in recognized (green curve) and unrecognized trials (orange curve) (low contrast study).* ***g*** *Target-locked pupil derivative modulation in*

correct (green curve) and in incorrect trials (orange curve) (backward masking study). ***h*** *Target-locked pupil derivative modulation in recognized (green curve) and unrecognized trials (orange curve) (low contrast study). The gray horizontal line represents the significant time points where the pupil size modulation or the pupil derivative modulation is significantly different for the two conditions ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. In all graphs, data are represented as mean ± standard error of the mean across participants.*

3.2.3.2 Analyses of pupil features used in the classifiers

For the analyses of the effect of pre-stimulus pupillary response in visual perception using classifiers, we extracted three different types of features: pupil size (percentage of the mean), relative pupil size (pupil dilation relative to the pre-cue baseline) and average pupil derivative. First, we studied how these were related to task performance in both studies and then we investigated if an SVM classifier was able to use them to predict the participant's performance.

3.2.3.2.1 Pupil size and derivative

Prior research suggests that participants with bigger pupil dilation may demonstrate a higher ability to detect stimuli and, consequently, better task performance than those with smaller dilation (3,43). In this way, our hypothesis was that a greater pupillary response was associated with better visual performance.

Contrary to our expectations, for the backward masking study, none of the measures showed statistically significant differences between the correct and incorrect trials (Figure 3.14 – a), Figure 3.14 – e) and Figure 3.14 – i)). Considering the pupil size, the results of the statistical analysis were: Shapiro-wilk test: $W = 0.881$, $p = 0.041$ and Wilcoxon test: $Z = -0.672$, $p = 0.501$. When we removed the baseline, the results were: Shapiro-wilk test: $W = 0.896$, $p = 0.064$ and paired-sample t-test: $t(15) = -0.341$, $p = 0.738$. Finally, when we consider the average pupil derivative the results were: Shapiro-wilk test: $W = 0.915$, $p = 0.138$ and paired-sample t-test: $t(15) = -0.105$, $p = 0.918$. In contrast to the results of the backward masking study, in the low contrast study, we can see significant differences in two of the three features analyzed. In both pupil size and average pupil derivative (Figure 3.14 – c) and Figure 3.14 – k)), the average value in the recognized trials is significantly higher than the average value in the unrecognized trials (pupil size - Shapiro-wilk test: $W = 0.944$, $p = 0.311$ and paired-sample t-test: $t(18) = 3.342$, $p = 0.004$; average pupil derivative - Shapiro-wilk test: $W = 0.965$, $p = 0.683$ and paired-sample t-test: $t(18) = 2.554$, $p = 0.020$). Conversely, the relative pupil size did not reveal any statistical difference between recognized and unrecognized trials (Shapiro-wilk test: $W = 0.970$, $p = 0.772$ and paired-sample t-test: $t(18) = 0.800$, $p = 0.434$). These results align with previous research findings (3,43).

Finally, in the backward masking study, we do not see any correlation between pupil features, such as relative pupil size and average pupil derivative, and RT (Figure 3.14 – f) and Figure 3.14 – j); relative pupil size - Shapiro-wilk test: $W = 0.942$, $p = 0.378$ and one-sample t-test: $t(15) = -0.256$, $p = 0.801$; average pupil derivative - Shapiro-wilk test: $W = 0.9012$, $p = 0.078$ and one-sample t-test: $t(15) = 0.086$, $p = 0.933$). However, in Figure 3.14 - b), we can see a significant positive correlation between pupil size and RT (Shapiro-wilk test: $W = 0.917$, $p = 0.132$ and one-sample t-test: $t(15) = 2.561$, $p = 0.022$). In the low contrast study, only the pupil size showed a significant negative correlation with RT (Figure 3.14 - d); Shapiro-wilk test: $W = 0.949$, $p = 0.392$ and one-sample t-test: $t(18) = -2.913$, $p = 0.009$). Since the RT is higher in unrecognized trials and the pupillary response is lower in these trials, we expected a negative correlation between these two measures. Thus, the obtained results for the feature pupil size aligns with the initially hypothesized expectation. The remaining measures, relative pupil size and average pupil derivative, did not show statistically significant correlations with reaction time (Figure 3.14 – h) and Figure 3.14 – l)) (relative pupil size - Shapiro-wilk test: $W = 0.876$, $p = 0.020$ and Wilcoxon test: $Z = -0.966$, $p = 0.334$; average pupil derivative - Shapiro-wilk test: $W = 0.926$, $p = 0.132$ and one-sample t-test: $t(18) = -1.477$, $p = 0.157$).

The results obtained in these analyses suggest that pupil-linked arousal fluctuations can influence the time needed to make the decision and probably the visual processing.

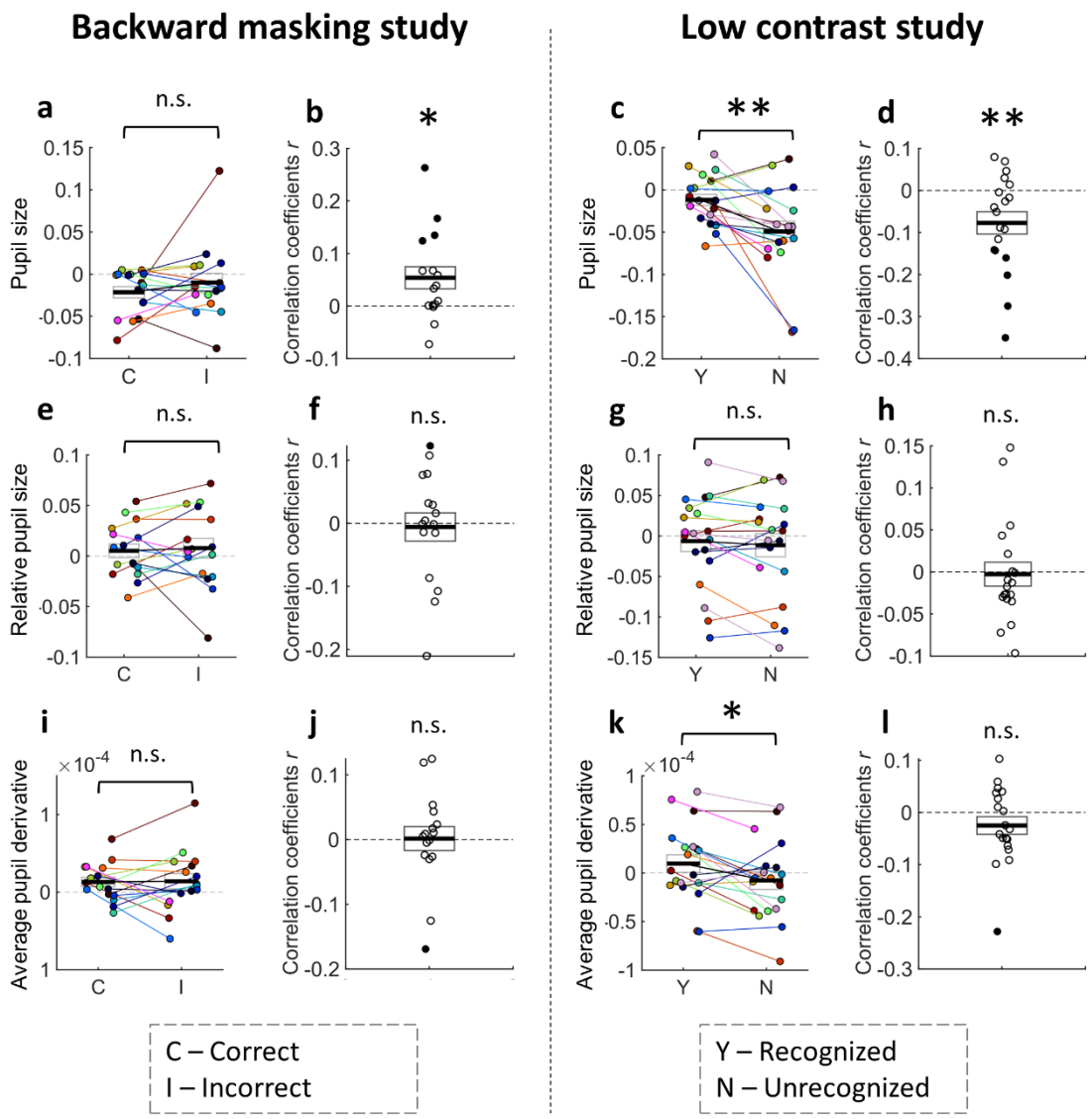


Figure 3.14 *– Fluctuations in pupil size measured in one second before target onset are associated with visual processing.* ***a*** *Relation between average pupil value and participant's accuracy (backward masking study).* ***b*** *Correlation between average pupil value and RT (backward masking study).* ***c*** *Relation between average pupil value and participant's recognition (low contrast study).* ***d*** *Correlation between average pupil value and RT (low contrast study).* ***e*** *Relation between relative pupil value and participant's accuracy (backward masking study).* ***f*** *Correlation between relative pupil value and RT (backward masking study).* ***g*** *Relation between relative pupil value and participant's recognition (low contrast study).* ***h*** *Correlation between relative pupil value and RT (low contrast study).* ***i*** *Relation between average pupil derivative and participant's accuracy (backward masking study).* ***j*** *Correlation between average pupil derivative and RT (backward masking study).* ***k*** *Relation between average pupil derivative and participant's recognition (low contrast study).* ***l*** *Correlation between average pupil derivative and RT (low contrast study). In sub-figures* ***a***, ***c***, ***e***, ***g***, ***i*** *and* ***k*** *the black line is the average pupil size/ average pupil derivative and the rectangle represents ± standard error of the mean. In sub-figures* ***b***, ***d***, ***f***, ***h***, ***j*** *and* ***l*** *the black line is the average correlation coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. Filled circles represent participants where correlation is statistically significant (p<0.05). *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.3.3 SVM: classification of visual performance based on pre-stimulus pupillary response

Then, we investigated whether there was information in the pupillary response that may facilitate the classifier's prediction of task performance. From Table 12, it is evident that, in the backward masking study, a maximum of 5% of the classifiers exhibit superior performance compared to random behavior, which means that the classifier was not able to extract activity patterns from these features. We believe that there exists a correlation between pupil size and its variation in relation to the cue (the results of the correlation analyses can be found in the appendix), which suggests that, when we combine these two measures, the patterns they deliver are similar (Figure 3.15 – a)). In the average pupil derivative, on the contrary, the performance is statistically lower compared to the performance observed when all measures are combined (Figure 3.15 – a)). In conclusion, in the backward masking study, as anticipated, it is not possible to predict task performance through pupillary response.

For the low contrast study, Table 12 shows that, in almost all combinations of measures, over 50% of the classifiers outperformed random behavior, except for the average pupil derivative. With average pupil derivative only 45% of the classifiers exhibited a behavior better than random. This result suggests that this parameter does not provide patterns of activity to the classifier. We then studied the influence of each measure individually. In Figure 3.15 – b), we can see that the performance of each measure separately is slightly inferior to the performance of all measures combined. Nonetheless, it is worth noting that the classifier's performance, when using pupil size and relative pupil size, does not exhibit statistically inferior results in comparison to its performance when incorporating all measures (Table 13). This result suggests that, in a certain way, the measures were correlated, so their combination did not provide very significant information to the classifier (results of correlation analyses are found in the appendix).

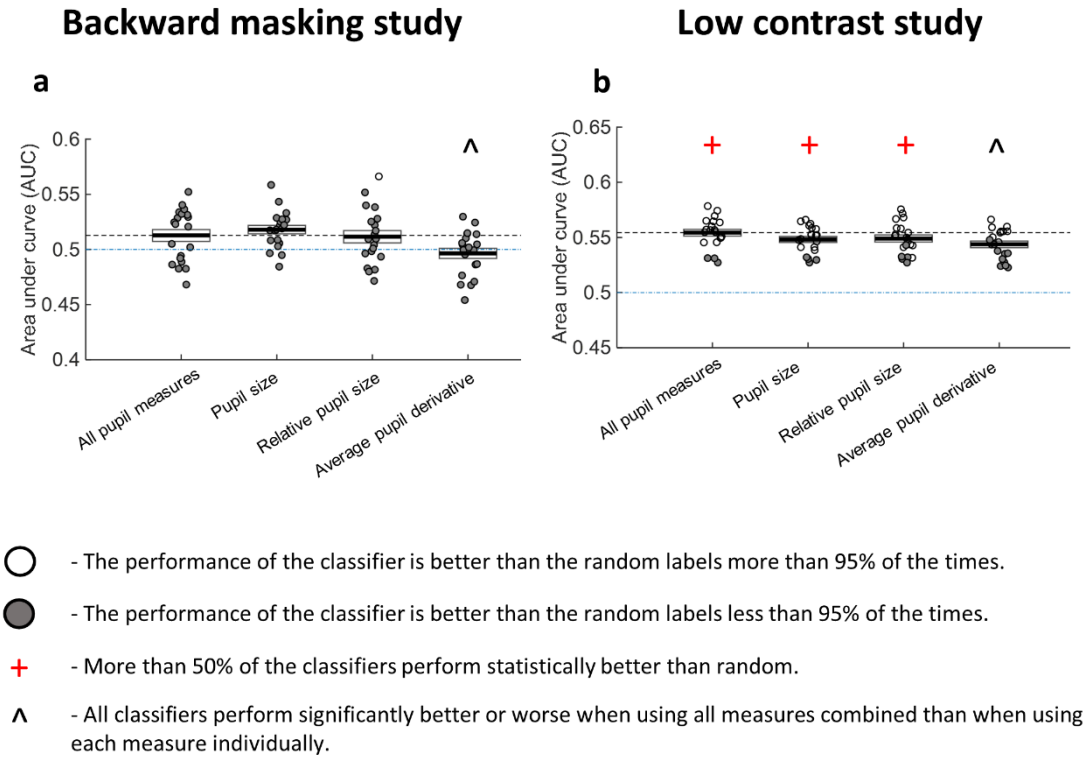
As mentioned, we did not expect the classifier to be able to predict accuracy in the backward masking study, but we expected the classifier to be able to predict stimulus recognition in the low contrast study. Thus, our results align with the anticipated expectations, suggesting that fluctuations in pupil size are associated with visual performance.

	Backward masking study	Low contrast study
All pupil measures	0%	**85%**
Pupil size	0%	**80%**
Relative pupil size	5%	**75%**
Average pupil derivative	0%	45%

Table 12 *– Percentage of classifiers using pupil measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	Low contrast study
Pupil size	$t\,(19) = -1.974$; $p = 0.246$	$t\,(19) = 1.913$; $p = 0.071$
Relative pupil size	$t\,(19) = -0.2025$; $p = 0.842$	$t\,(19) = 1.4858$; $p = 0.154$
Average pupil derivative	**$t\,(19) = 2.5401$; $p = 0.020$**	**$t\,(19) = 2.634$; $p = 0.016$**

Table 13 *- Comparison between the use of each pupil measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all pupil measures were combined. The statistically significant analyses are highlighted.*



○ - The performance of the classifier is better than the random labels more than 95% of the times.

● - The performance of the classifier is better than the random labels less than 95% of the times.

\+ - More than 50% of the classifiers perform statistically better than random.

^ - All classifiers perform significantly better or worse when using all measures combined than when using each measure individually.

Figure 3.15 *- Pre-stimulus pupillary activity allows predicting the participant's recognition in the low contrast study but does not allow predicting task performance in the backward masking study.* ***a*** *AUC using a SVM model in the backward masking study.* ***b*** *AUC using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all pupil measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed. +: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random. ^: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

3.2.4 Effect of pre-stimulus blinking activity on visual perception

3.2.4.1 Study of the time course of blinking rate variation evoked by the warning cue

The blink rate is thought to be a measure of cognitive load and task engagement. Subjects tend to blink less in demanding cognitive tasks, which reflects a higher level of attentional focus and cognitive effort (48,49). Based on this, we expected that the alerting stimulus would induce a reduction in the blinking rate. In fact, in Figure 3.16 – a) and Figure 3.16 – b), we can observe a significant reduction in the blinking rate. However, the reduction occurred only about one second after the warning cue onset. In Figure 3.16 – a) and Figure 3.16 – b), we can see an abrupt increase in the blinking rate after the cue onset. When we get an alarm signal or any abrupt or unexpected stimulus, our body automatically reacts to protect the eye, which may cause a blink. Therefore, our results are in line with expectations. That said, we consider that the changes produced by the state of expectation are verified only from the second one.

About one second after the visual stimulus is presented, there is a significant increase in blinks rate, coinciding with the point at which the subject made a choice and provided a response with the chosen category option (in the task performance section, we can see that the average reaction time in both studies is approximately one second after stimulus onset - Figure 3.2). We believe this is a restoration of normality, causing an increase in blink activity.

In the backward masking study, when comparing correct with incorrect trials, we see a significant difference in blinks' rate; however, this difference occurred only one second after visual stimulus onset (Figure 3.16 – g)). In turn, in the low contrast study, we do not see any significant difference when comparing recognized vs unrecognized trials (Figure 3.16 – h)). As blinks can interrupt the flow of visual information, it is expected that closer to the target, the blinks rate would be higher in incorrect trials and in unrecognized trials. If a participant blinks at the moment of the stimulus, it can result in a loss of information, which can lead to the

participant not detecting the stimulus and mistaking the stimulus category (46). In Figure 3.16 and Figure 3.16 – h); however, we cannot see significant differences between the trial types in both studies.

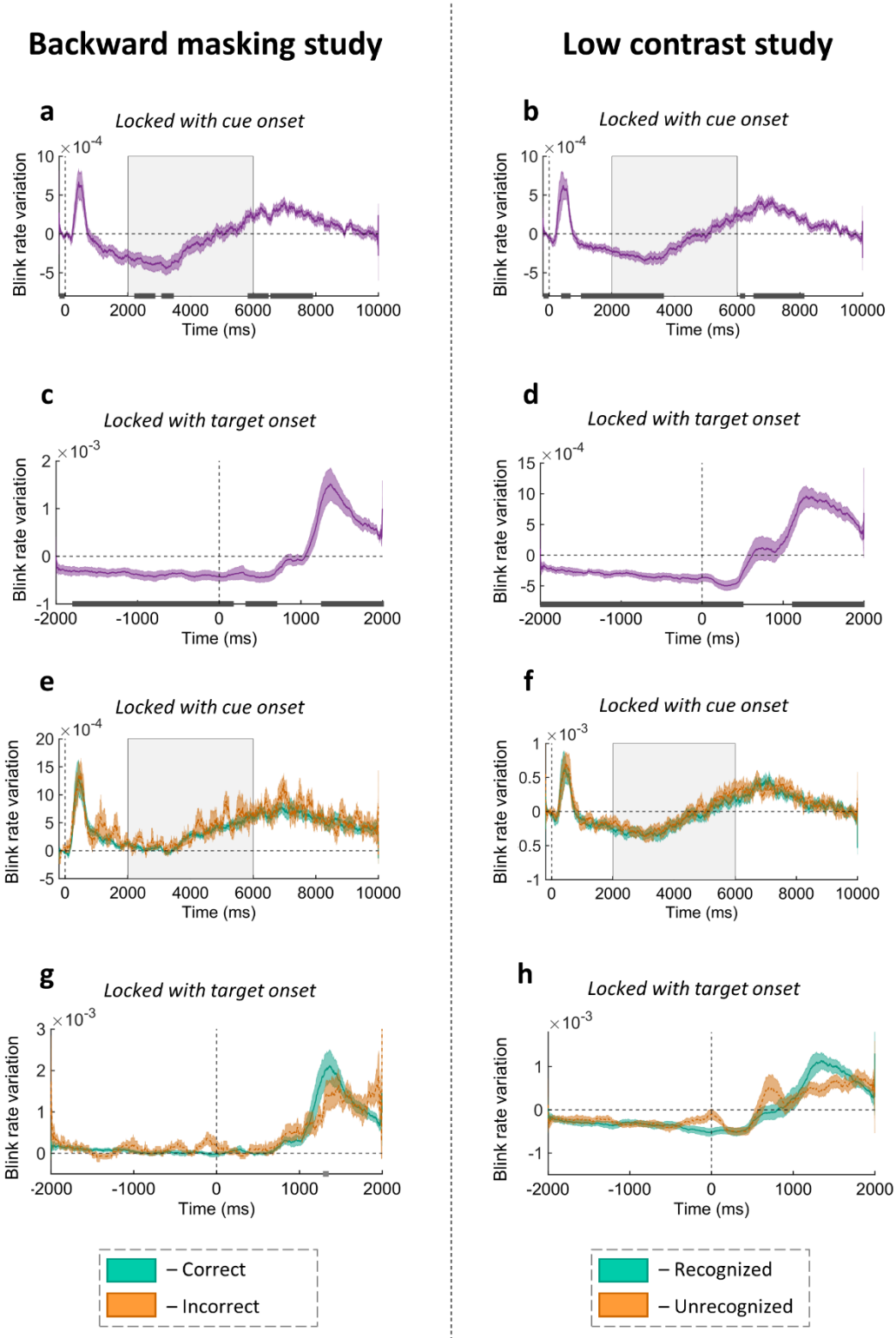


Figure 3.16 *– The cue induced a decrease in the blink rate.* ***a and b*** *Cue-locked blink rate modulation (a - backward masking study; b - low contrast study).* ***c and d*** *Target-locked blink rate modulation (a - backward masking study; b - low contrast study). The gray horizontal line represents the significant time points where*

the blink rate is significantly different from zero ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. ***e*** *Cue-locked blink rate modulation in correct trials (green curve) and incorrect trials (orange curve) (backward masking study).* ***f*** *Cue-locked blink rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study).* ***g*** *Target-locked blink rate modulation in correct trials (green curve) and incorrect trials (orange curve) (backward masking study).* ***h*** *Target-locked blink rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study). The gray horizontal line represents the significant time points where the blink rate is significantly different for the two conditions ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. In all graphs, data are represented as mean ± standard error of the mean across participants.*

3.2.4.2 Analyses of blinking activity features used in the classifiers

For the analyses of the effect of pre-stimulus blinking activity in visual perception using classifiers, we extracted two different types of features: blink rate and the temporal distance between the last blink and the visual stimulus onset. First, we studied how these were related to task performance in both studies and then we investigated if an SVM classifier was able to use them to predict the participant's performance.

3.2.4.2.1 Blink rate and temporal distance between the last blink and the visual stimulus onset

As seen in Figure 3.16 – g) and Figure 3.16 – h), the blink rate closer to stimulus presentation is higher in incorrect trials, as well as in unrecognized trials, however, this difference was not statistically significant. To further study this effect, we compared the blink rate measured in an interval between one second after cue onset and stimulus onset in incorrect trials against the blink rate in correct trials, as well as between recognized and unrecognized trials.

In both the backward masking study and the low contrast study (Figure 3.17 – e) and Figure 3.17 – g)), we see no significant differences in blink rate between trial types in both studies (backward masking study: Shapiro-wilk: $W = 0.848$, $p = 0.018$ and Wilcon test: $Z = 0.950$ $p = 0.330$; low contrast study: Shapiro-wilk test: $W = 0.913$, $p = 0.078$ and paired-sample t-test: $t(18) = -1.976$, $p = 0.064$).

In addition to studying the blink rate, we also studied the temporal distance between the last blink and the target onset. Once again, the differences verified between the trial types turned out not to be statistically significant (backward masking study (Figure 3.17 – a)): Shapiro-wilk test: $W = 0.938$, $p = 0.305$ and paired-sample t-test: $t(14) = 2.002$, $p = 0.065$; low contrast study (Figure 3.17 – c)): Shapiro-wilk test: $W = 0.983$, $p = 0.974$ and paired-sample t-test: $t(18) = 0.852$, $p = 0.405$).

It would be expected that the blinking activity would influence stimulus detection, since if the participants blink too close to the stimulus, it might affect stimulus detection. However, this is not what happens. In fact, blinks are rare events whose frequency is significantly reduced in the state of expectation. As a result, the probability of a blink occurring near to the stimulus becomes very low, thereby only influencing sporadically visual performance.

Moreover, we then investigated the relationship between RT and blinking features. From Figure 3.17, we can see that none of the features produced significant correlations.

Considering the relation between RT and participant's visual performance and considering that the blink rate is higher in incorrect and in unrecognized trials, we expected to see a positive correlation between the RT and the blink rate. In fact, in both studies, we observe an average positive value of the correlation values, however these were not significantly different from zero (backward masking study (Figure 3.17 – f)): Shapiro-wilk test: $W = 0.952$, $p = 0.475$ and one-sample t-test: $t(14) = 1.045$, $p = 0.314$; low contrast study (Figure 3.17 – h)): Shapiro-wilk test: $W = 0.983$, $p = 0.970$ and one-sample t-test: $t(18) = 2.082$, $p = 0.052$).

Similarly to the analysis conducted for blink rate, we also examined the relationship between the temporal distance between last blink and visual stimulus onset and RT. When we considered the temporal distance, we expected a negative correlation, since in incorrect and unrecognized trials the temporal distance is smaller and the RT is higher. However, the correlation values, once again, were not statistically different from zero (backward masking study (Figure 3.17 – b)): Shapiro-wilk test: $W = 0.827$, $p = 0.010$ and Wilcoxon test: $Z = 1.010$, $p = 0.188$; low contrast study (Figure 3.17 – d)): Shapiro-wilk test: $W = 0.971$, $p = 0.799$ and one-sample t-test: $t(18) = -0.250$, $p = 0.806$).

The results obtained suggest that, overall, there seems to be no significant relationship between blinking activity and the time required to make the decision. These findings are consistent with previous observations and indicate that, probably, blinking activity does not significantly influence visual processing.

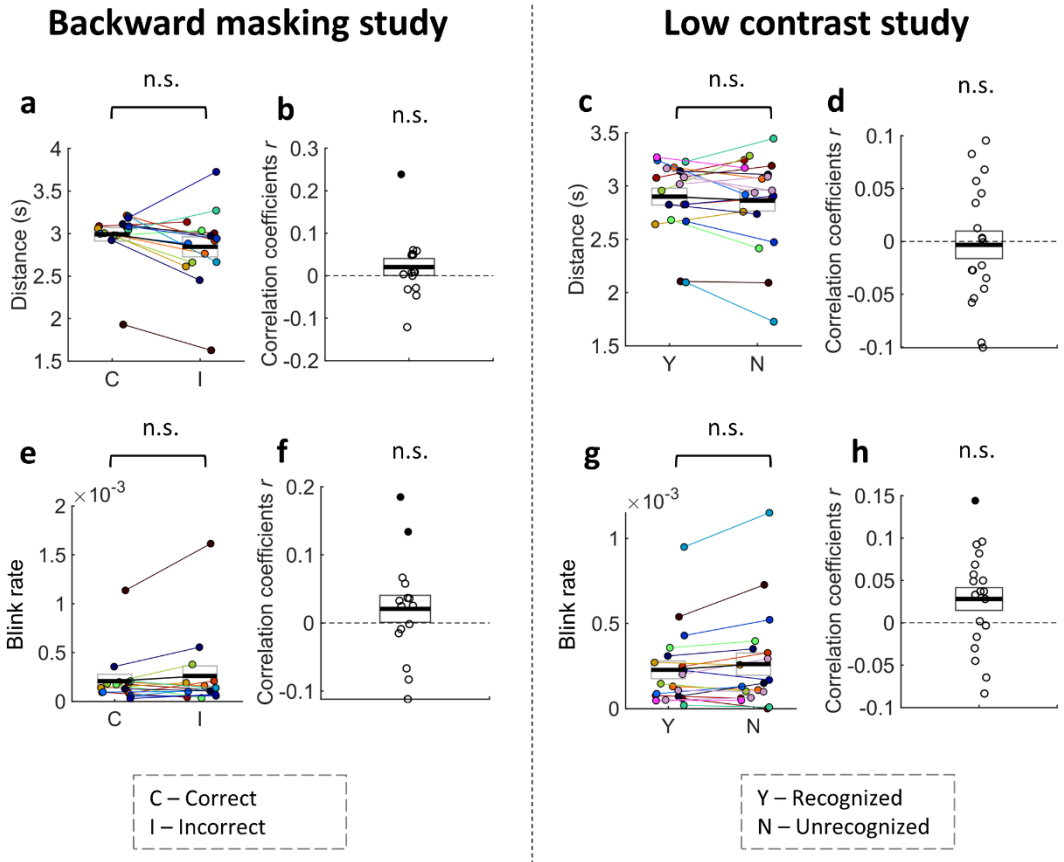


Figure 3.17 *– Blinking activity, measured in an interval between one second after cue onset and stimulus onset, does not influence visual performance.* ***a*** *Relation between temporal distance of the last blink and participant's accuracy (backward masking study).* ***b*** *Correlation between temporal distance of the last blink and RT (backward masking study).* ***c*** *Relation between temporal distance of the last blink and participant's recognition (low contrast study).* ***d*** *Correlation between temporal distance of the last blink and RT (low contrast study).* ***e*** *Relation between blink rate and participant's accuracy (backward masking study).* ***f*** *Correlation between blink rate and RT (backward masking study).* ***g*** *Relation between blink rate and participant's recognition (low contrast study).* ***h*** *Correlation between blink rate and RT (low contrast study). In sub-figures* ***a***, ***c***, ***e*** *and* ***g*** *the black horizontal line is the average temporal distance of the last blink/blink rate, and the rectangle represents ± standard error of the mean. In sub-figures* ***b***, ***d***, ***f*** *and* ***h*** *the black horizontal line is the average correlation coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. Filled circles represent participants where correlation is statistically significant (p<0.05). *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.4.3 SVM: classification of visual performance based on pre-stimulus blinking activity

Table 14 demonstrates that, using the blinking parameters from the backward masking study, and employing a permutation test, a maximum of 15% of the classifiers exhibit superior performance when compared to random behavior. Analyzing Figure 3.18 – a), we can see that, when we used only the blink rate as input, the performance obtained greatly surpasses that achieved by other combinations of measures. However, even though this parameter has an approximate average performance of 0.65, when the permutation

tests were applied, only 5% of the classifiers presented a behavior superior to randomness. These results suggest that this average performance may have occurred by chance. Then, we studied the contribution of each measure to the classifier. Our initial expectation was that the combination of all measures would yield classifiers with statistically superior performance compared to classifiers using each measure individually. However, this hypothesis is not supported by the results (Figure 3.18 – a)). Considering the average performance achieved when using only the blink rate as input, it would be expected that this measure would provide important and extra information to the classifier, thereby favoring its predictive capacity. However, we found that the combination of this measure with the distance of the last blink within each trial did not produce improvements in performance. This result aligns with the previous conclusions, where it is suggested that the performance achieved by using blink rate as a measure was a consequence of chance. In other words, if this measure effectively contained pertinent information, the classifier would have been able to extract activity patterns that would allow distinguishing between correct and incorrect trials.

The same analyses were performed for the low contrast study. From Table 14 we can see that, as observed for the backward masking study, the classifiers cannot predict the participant's recognition through the blinking activity, since only 45% of the classifiers present a superior performance than the random one. When combining the different measures of blinking activity, we expected a statistically superior performance compared to classifiers' performance using each measure individually. However, once again, this did not happen (Figure 3.18 – b)). In Figure 3.18 – b), we can verify that the average performance obtained through the combination of all measures is practically equivalent to the average performance when each measure was used individually. We believe that this behavior may be attributed to a possible correlation between the temporal distance of the last blink and blink rate (the results of the correlation analyses are in the appendix). This correlation it is logical, because if a person blinks a lot, there will be a probability of blinking very close to the target.

In conclusion, the classifier was not able to predict participant's visual performance using blinking activity parameters as features.

	Backward masking study	Low contrast study
All blink measures	15%	45%
Blink rate	5%	5%
Distance of the last blink	5%	45%

Table 14 *- Percentage of classifiers using blinking features as input that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that a minimum of 50% of the classifier exhibits a performance surpassing random chance.*

	Backward masking study	Low contrast study
Blink rate	***t (19)* = -15.283; *p* <0.001**	*t (19)* = -1.345; *p* = 0.194
Distance of the last blink	*t (19)* = 0.4592; *p* = 0.651	*t (19)* = -0.250; *p* = 0.806

Table 15 *- Comparison between the use of each blink measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all blinking measures were combined. The statistically significant analyses are highlighted.*

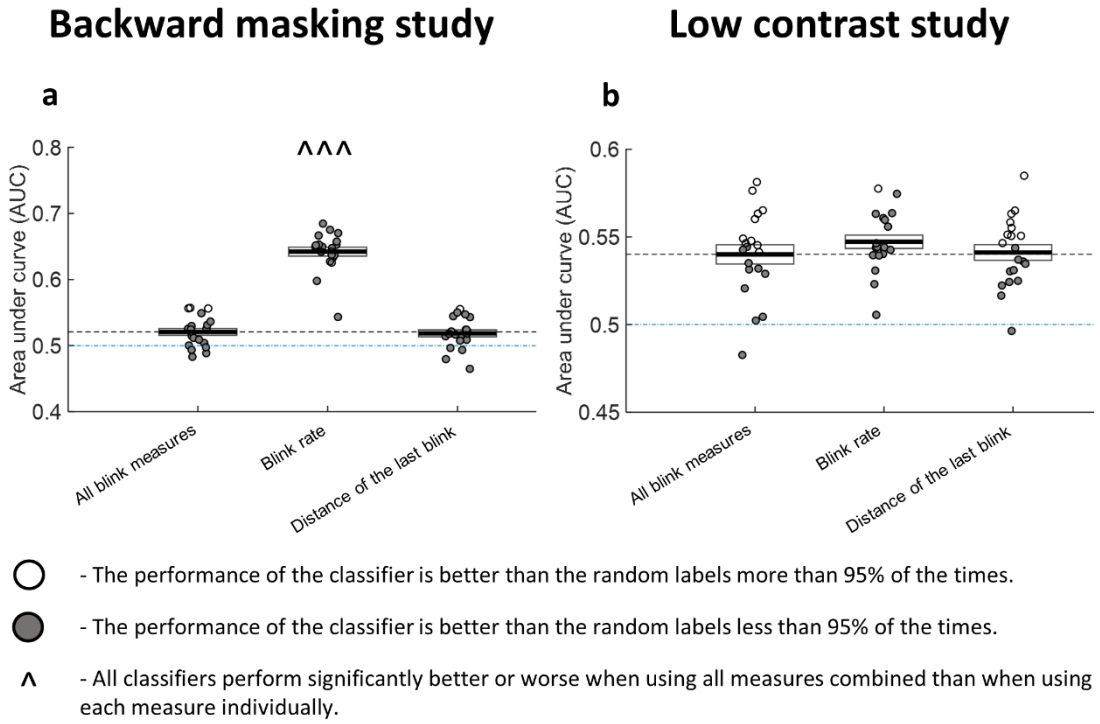


Figure 3.18 *- Pre-stimulus blinking activity does not allow predicting task performance or participants recognition.* ***a*** *AUC using a SVM model in the backward masking study.* ***b*** *AUC using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all blink measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed. ^:*

When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.

3.2.5 Effect of pre-stimulus saccadic activity on visual perception

3.2.5.1 Study of the time course of saccades rate variation evoked by the warning cue

In the state of expectation, the individual tends to momentarily stop all their eye movements, namely the saccades (7–9). The saccades are, among other things, responsible for shifting the focus of the fovea, enabling us to efficiently and quickly locate crucial information (53). However, in some tasks, saccades can affect vision and introduce errors (47,52).

In Figure 3.19 – a) and Figure 3.19 – b), we can see that, shortly after the warning cue onset, there is a rapid and significant reduction in saccadic activity followed by a rapid and significant increase in its rate. This behavior can be explained by the eye looking for a possible threat (suggested by the warning cue), which can cause a saccade. As in the blinks section, we consider that the changes produced by the state of expectation are verified only from the second one.

In Figure 3.19 – c) and Figure 3.19 - d), it is possible to see that after stimulus presentation there is a significant increase in saccadic activity associated with the appearance of the response prompt that evokes eye movements.

Again, we also analyzed the differences in the saccades rate in correct and incorrect trials, as well in recognized and unrecognized trials. As mentioned, saccades can impair vision due to the rapid movement of the image, which can cause blurriness and loss of important information. In this way, it is expected that close to the target, the saccades rate will be higher for incorrect and unrecognized trials. This behavior is evident in Figure 3.19 – h), where it can be observed that, at 200 ms before stimulus onset, there is a significant difference in saccades rate between recognized and unrecognized trials in the low contrast study. We do not see any significant difference when comparing correct vs incorrect trials in the backward masking study.

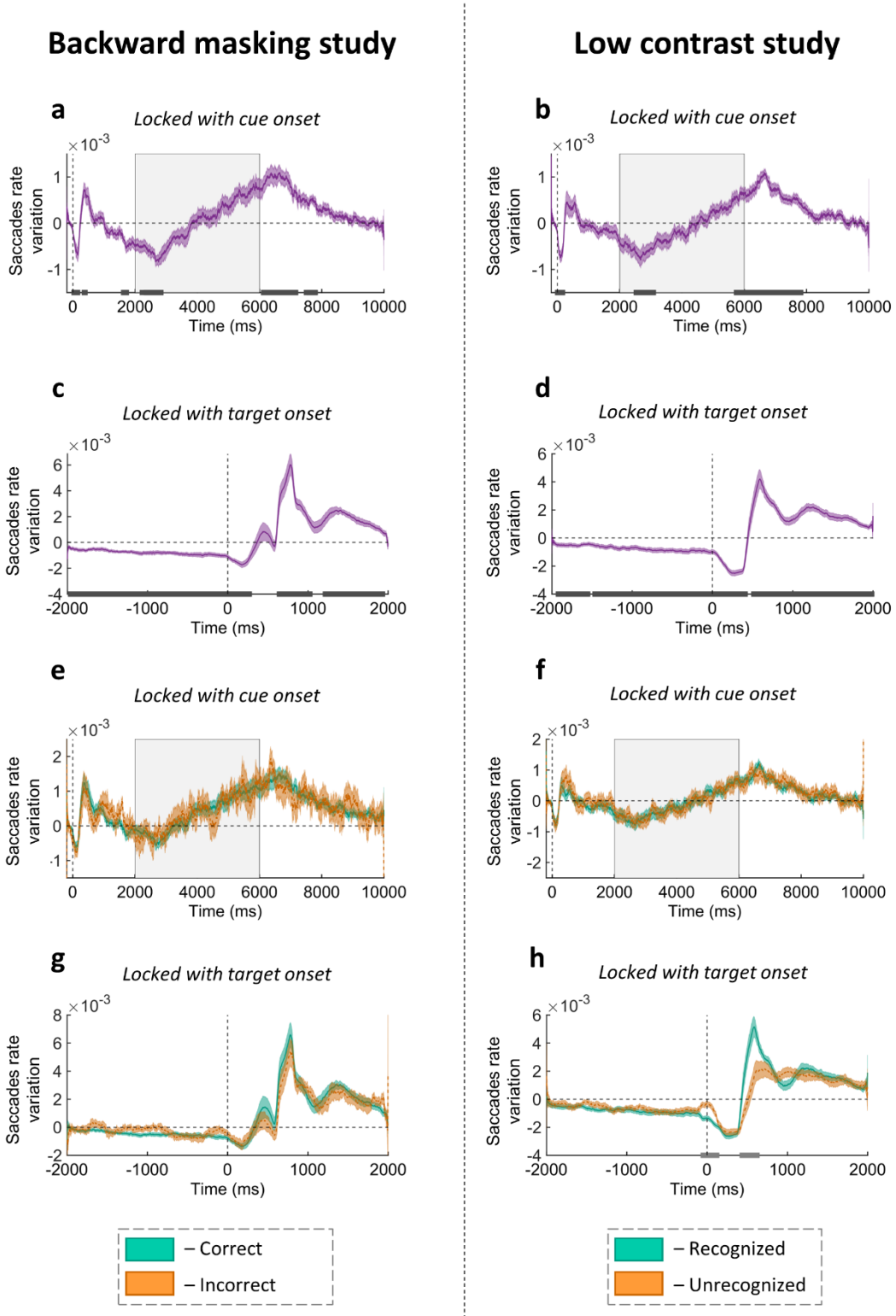


Figure 3.19 *– The cue induced a decrease in the saccades rate.* ***a and b*** *Cue-locked saccades rate modulation (a - backward masking study; b - low contrast study).* ***c and d*** *Target-locked saccades rate modulation (a - backward masking study; b - low contrast study). The gray horizontal line represents the significant time points where the saccades rate is significantly different from zero ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented.* ***e*** *Cue-locked saccades rate modulation in correct trials (green curve) and incorrect trials (orange curve) (backward masking study).* ***f*** *Cue-locked saccades rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study).* ***g*** *Target-locked saccades rate modulation in correct trials (green curve) and incorrect trials (orange curve)*

(backward masking study). ***h*** *Target-locked saccades rate modulation in recognized trials (green curve) and unrecognized trials (orange curve) (low contrast study). The gray horizontal line represents the significant time points where the saccades rate is significantly different for the two conditions ($p < 0.05$). The gray rectangle represents the time window in which the stimulus is presented. In all graphs, data are represented as mean ± standard error of the mean across participants.*

3.2.5.2 Analyses of saccadic activity features used in the classifiers

For the analyses of the effect of pre-stimulus saccadic activity in visual perception using classifiers, we extracted two different types of features: average saccades rate within the interval between one second after cue onset and visual stimulus onset and the temporal distance between the last saccade and the visual stimulus onset. First, we studied how these were related to task performance in both studies and then we investigated if an SVM classifier was able to use them to predict the participant's performance.

3.2.5.2.1 Saccades rate and temporal distance between the last saccade and the visual stimulus onset

For the study of saccades rate, we expected that the higher the rate, the more likely the participant would not detect the stimulus and thus miss the category. In fact, this was the observed behavior in the low contrast study, where we can see significant differences in saccades rate between the trial types (Figure 3.20 – g); Shapiro-wilk test: $W = 0.945$, $p = 0.270$ and paired-sample t-test: $t(18) = -2.842$, $p = 0.011$). In turn, the backward masking study revealed no significant differences in saccades rate between correct and incorrect trials (Figure 3.20 – e); Shapiro-wilk test: $W = 0.939$, $p = 0.283$ and paired-sample t-test: $t(15) = -1.985$, $p = 0.066$).

Then, we studied the temporal distance between the last saccade and the target presentation. In both the backward masking study and the low contrast study, we see significant differences in the temporal distance of the last saccade between the trial types (correct/incorrect trials in the backward masking study and recognized/unrecognized trials in the low contrast study) (backward masking study (Figure 3.20 – a)):Shapiro-wilk test: $W = 0.938$, $p = 0.274$ and paired-sample t-test: $t(15) = -1.985$, $p = 0.042$; low contrast study (Figure 3.20 – c)): Shapiro-wilk test: $W = 0.970$, $p = 0.773$ and paired-sample t-test: $t(18) = 2.381$, $p = 0.028$). These findings are coherent, since a smaller temporal distance between the last saccade and the target leads to a saccade occurring at the moment of stimulus onset, which can impair vision.

The results obtained suggest that pre-stimulus saccadic activity is associated with visual stimulus recognition.

Moreover, we then investigated the relationship between RT and saccadic features. From Figure 3.20, we can see that none of the features produced significant correlations (correlation between RT and saccades rate - backward masking study (Figure 3.20 – f)): Shapiro-wilk test: W = 0.881, p = 0.040 and Wilcoxon test: Z = 1.913, p = 0.059; low contrast study (Figure 3.20 – h)): Shapiro-wilk test: W = 0.933, p = 0.198 and one-sample t-test: $t(18)$ = 0.042, p = 0.967; correlation between RT and the temporal distance between last saccade and visual stimulus onset - backward masking study (Figure 3.20 – b)): Shapiro-wilk test: W = 0.963, p = 0.720 and one-sample t-test: $t(15)$ = -1.545, p = 0.143; low contrast study (Figure 3.20 – d)): Shapiro-wilk test: W = 0.945, p = 0.326 and one-sample t-test: $t(18)$ = 0.119, p = 0.907).

The results obtained in both studies suggest that saccadic activity might not influence the time needed to make the decision, i.e., the reaction time. Still, according to the first analyses, there seems to be a correlation between saccadic activity and visual processing.

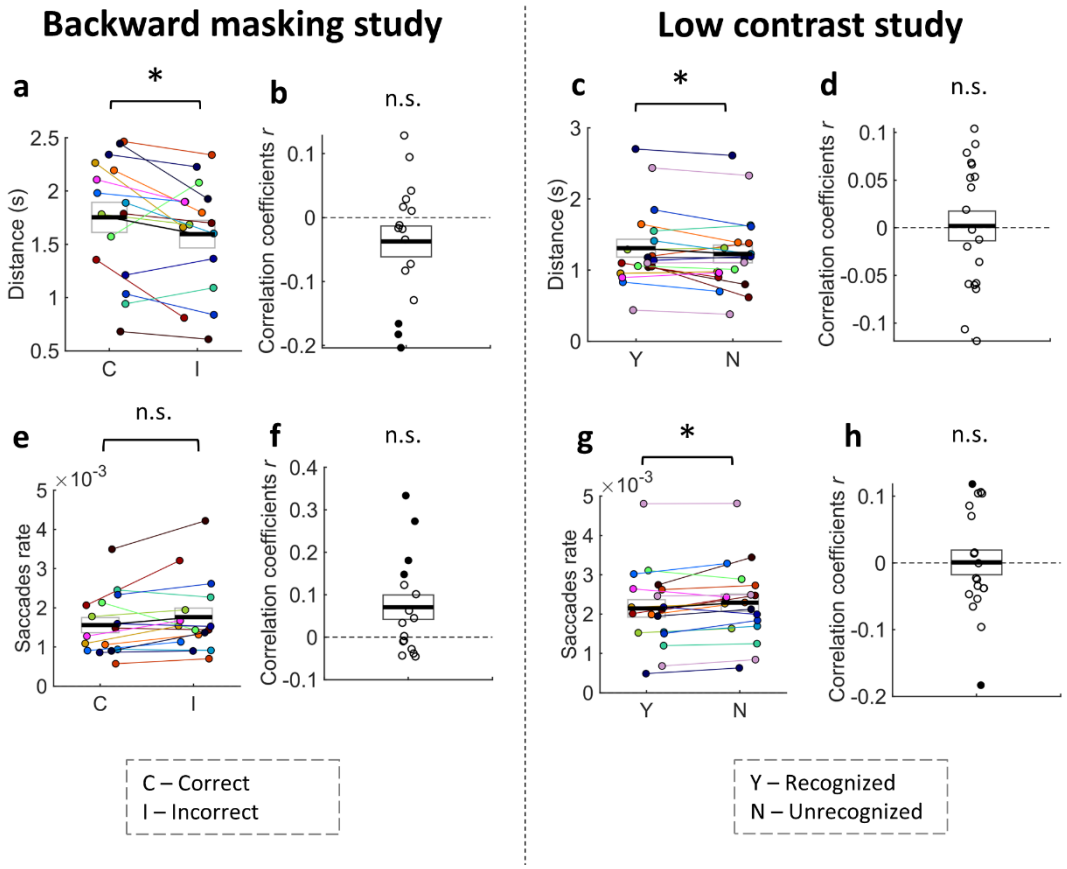


Figure 3.20 *– The saccadic activity, measured in an interval between one second after cue onset and stimulus onset, is associated with visual performance.* ***a*** *Relation between temporal distance of the last saccade and participant's accuracy (backward masking study).* ***b*** *Correlation between temporal distance of the last saccade and RT (backward masking study).* ***c*** *Relation between temporal distance of the last saccade and participant's recognition (low contrast study).* ***d*** *Correlation between temporal distance of the last saccade and RT (low contrast study).* ***e*** *Relation between saccades rate and participant's accuracy (backward masking study).* ***b*** *Correlation between saccades rate and RT (backward masking study).* ***c*** *Relation between saccades rate and participant's recognition (low contrast study).* ***d*** *Correlation between*

*saccades rate and RT (low contrast study). In sub-figures **a**, **c**, **e** and **g** the black line is the average temporal distance of the last saccade/saccades rate and the rectangle represents ± standard error of the mean. In sub-figures **b**, **d**, **f** and **h** the black line is the average correlation coefficients and the rectangle represents ± standard error of the mean. Individual circles represent data from each participant. Filled circles represent participants where correlation is statistically significant (p<0.05). *p<0.05, **p≤0.01, ***p≤0.001, n.s.: not significant.*

3.2.5.3 SVM: classification of visual performance based on pre-stimulus saccadic activity

In the backward masking study, Table 16 demonstrates that only a maximum of 15% of the classifiers demonstrate superior performance compared to random behavior., i.e., in the backward masking study the classifiers were not able to predict visual performance with saccadic features. Similar to the observations made with blinking activity, we find a parallel pattern here: when using only saccades rate as input, the performance obtained is significantly higher compared to the one achieved by other combinations of measures. Again, through the application of a permutation test, we verify that only 5% of the classifiers exhibit a behavior statistically superior to random. Once again, these results suggest that the performance using only saccades rate is not statistically reliable and may have occurred due to random chance. Subsequently, we studied the individual contribution of each measure to the classifier. Initially, the expectation was that the combination of all measures would yield classifiers with statistically superior performance compared to classifiers using each measure individually. However, this hypothesis is not corroborated by the results (Figure 3.21 – a)). Considering the high performance achieved when using only the saccades rate as input, it was presumed that the saccades rate would provide significant supplementary information to the classifier, thus augmenting its predictive capability. Nevertheless, again, the combination of this measure with the distance of the last saccade within each trial did not produce improvements in classifier's performance. This result aligns with the earlier conclusions, where it is suggested that the augmented performance achieved through the utilization of saccades rate as a measure was more likely a consequence of chance.

In the low contrast study, from Table 16, we can see that, as observed for the backward masking study, the classifiers were not capable of predicting visual performance based on saccadic activity, given that only 15% of the classifiers present a superior performance compared to randomness. In this present study, we observed a possible correlation between the temporal distance of the last saccade and saccades rate (the results of the correlation analyses are in the appendix). This hypothesis is supported by Figure 3.21 - b). The figure illustrates the performance of classifiers employing each measure individually was nearly identical to the performance achieved through the combination of both measures. This outcome turns out to make sense, because if a participant has a higher rate of saccades in the

considered interval, it is possible that one of the saccades occurred very close to the stimulus, thus influencing the temporal distance of the last saccade to the stimulus. Thus, the combination of these two measures does not provide additional information to the classifier.

Contrary to our expectations, the classifier was not able to predict visual performance in both studies. We consider that these results may be attributed to a potential bias in the conducted analyses. Let us consider that in two trials the participant had a saccade very close to the stimulus, which impaired the visualization and recognition of the stimulus. Consequently, when analyzing the temporal distance from the last saccade to the stimulus, the distance in these specific trials would be minimal. As a result, even a minimal distance in these two trials could influence the overall average behavior. Thus, what initially appeared to be a different behavior between the conditions under study in both studies could have resulted only from the behavior observed in these two trials. In these particular trials (used purely as an example) recognizable patterns were observed, while the remaining incorrect and unrecognized trials did not exhibit patterns in saccadic activity. That said, the classifier can be capable of recognizing the patterns in the mentioned trials, but it is not able to identify patterns in the remaining trials, as such patterns do not exist. In conclusion, pre-stimulus saccadic activity did not allow us to predict stimulus recognition using the SVM algorithm.

	Backward masking study	**Low contrast study**
All saccade measures	10%	15%
Saccades rate	5%	5%
Distance of the last saccade	15%	5%

Table 16 - *Percentage of classifiers using saccadic features as input that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that a minimum of 50% of the classifier exhibits a performance surpassing random chance.*

	Backward masking study	**Low contrast study**
Saccades rate	***t (19)* = -5.572; *p* <0.001**	*t (19)* = -0.099; *p* = 0.922
Distance of the last saccade	*t (19)* = 1.074; *p* = 0.296	*t (19)* = 0.634; *p* = 0.534

Table 17 *- Comparison between the use of each saccade measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all saccades' measures were combined. The statistically significant analyses are highlighted.*

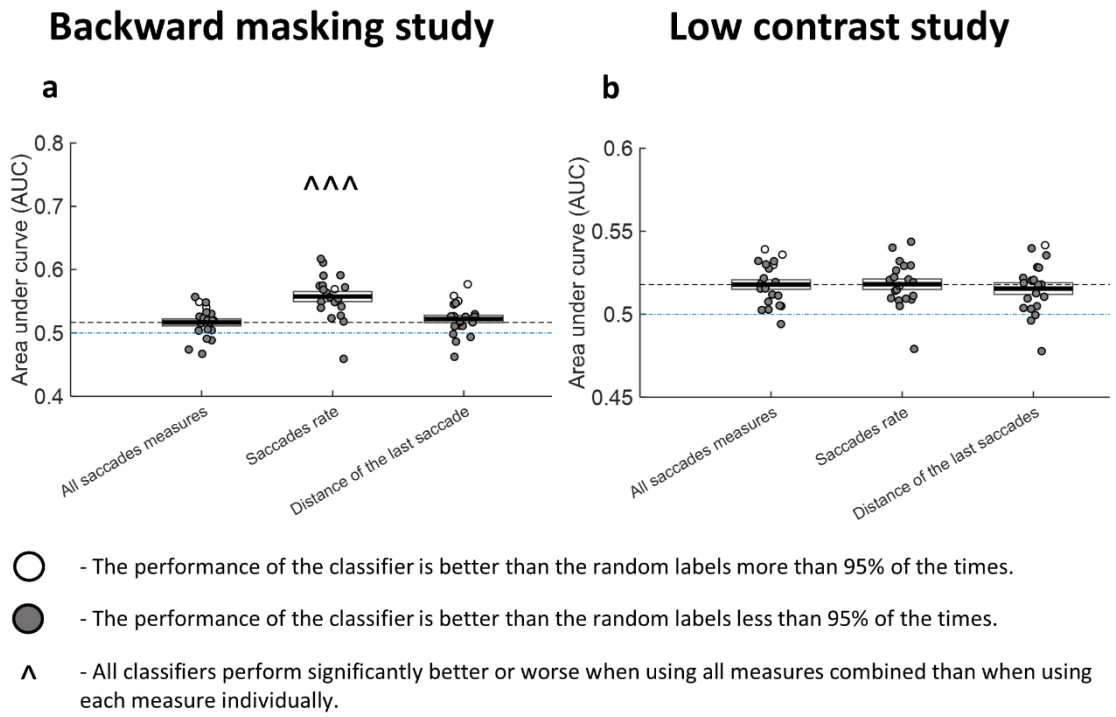


Figure 3.21 - *Pre-stimulus saccadic activity does not allow predicting task performance or participants recognition.* ***a*** *AUC using a SVM model in the backward masking study.* ***b*** *AUC using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all saccade measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***^****: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

3.2.6 Effect of pre-stimulus neural activity on visual perception

3.2.6.1 Study of the time course of EEG activity variation evoked by the warning cue and extraction of EEG features

During the state of attentive anticipation, a slow negative shift in brain electrical potential, the CNV, can be observed in the EEG (67,68). This wave reflects the preparatory potential evoked by the warning stimulus. The CNV is typically observed in the fronto-central region of the scalp and is thought to originate in the anterior cingulate cortex which is located in the medial frontal cortex (129). The CNV shows maximum amplitude in the frontocentral channel FCz and, therefore, we chose this channel for analyses (Figure 3.22 - a) and Figure 3.22 – b)).

Since, the evoked negative waveform reflects the preparation to the upcoming stimulus, we expected that a more negative amplitude of the EEG signal was related to a higher capacity to perceive the stimuli. In Figure 3.22 – c) and Figure 3.22 - d); however, no significant differences were observed between trial types. Nonetheless, we observed significant differences in the ERPs evoked by the recognized and the unrecognized stimuli, after target onset, in the low contrast study. These differences suggest that the neural processing is different in these two conditions in the time window between 250 to 500 ms after cue onset, possibly representing a response that indicates conscious processing of the stimuli (Figure 3.22 – d)). In contrast, backward masking study did not exhibit such significant differences (Figure 3.22 – c)).

It should be noted that the presented signal corresponds only to the data extracted from the FCz channel. In order to analyze all the channels, we used a different approach. For the analyses of the effect of pre-stimulus neural activity in visual perception, we extracted the average EEG activity from each electrode one second before the stimulus presentation. Then, we segregated the data based on the participants' responses. In both Figure 3.22 – e) and Figure 3.22 – f), we found no significant differences for any of the channels when comparing across conditions.

In none of the analyses there seems to be an association between the pre-stimulus neural activity and the visual processing. These results contradict the findings of (5,6), which suggest that the internal state of the brain contributes to the stimuli perception. We believe that this discrepancy in the results may be associated with the low spatial resolution of the EEG, as well as its susceptibility to significant noise interference.

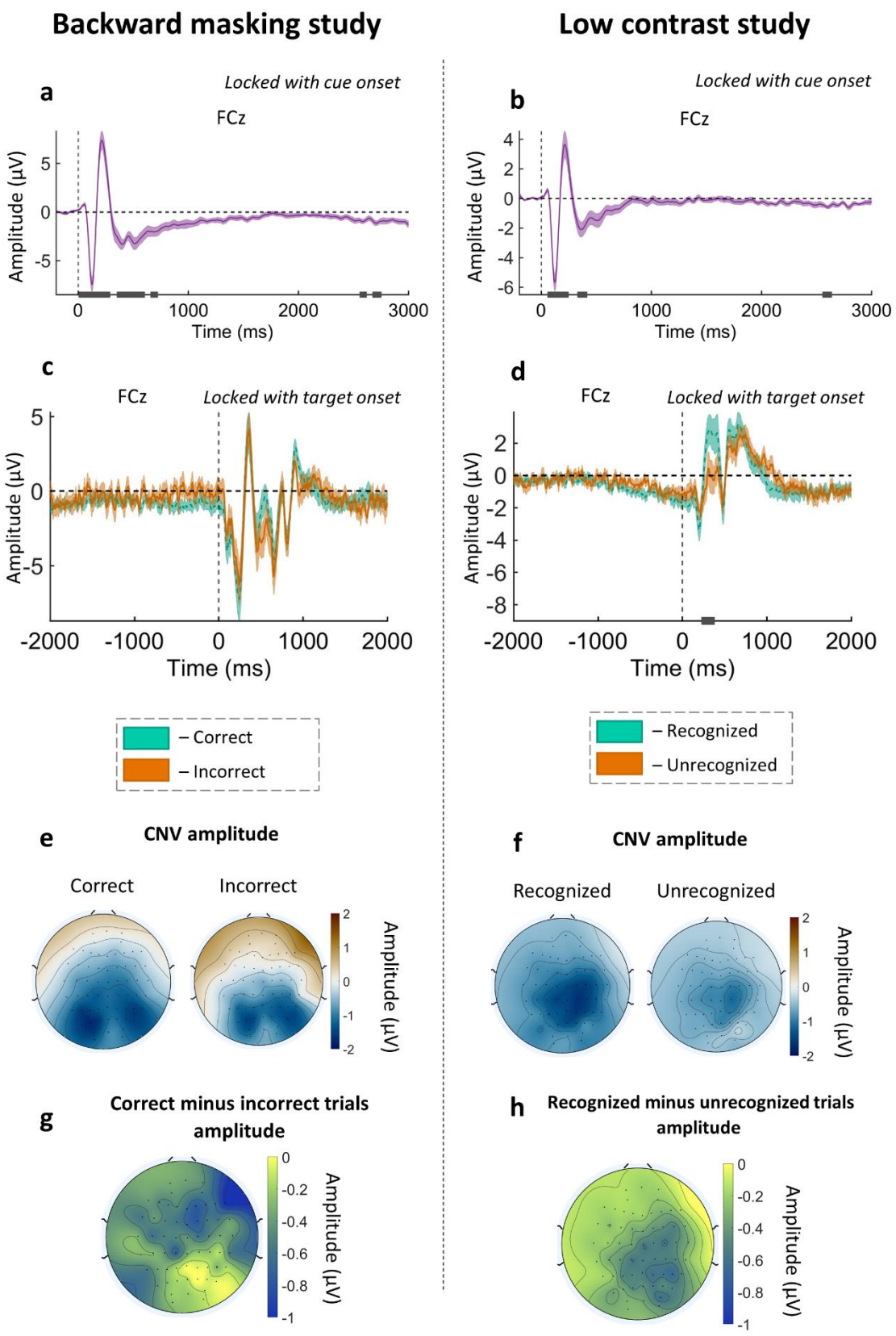


Figure 3.22 *– Brain pre-stimulus activity is not significantly associated with visual performance.* ***a and b*** *CNV induced by the cue (a - backward masking study; b - low contrast study). The gray horizontal line represents the time windows where the EEG signal is significantly different from zero ($p < 0.05$).* ***c*** *EEG activity in correct (green line) and in incorrect trials (orange line) locked with target onset (backward*

masking study). ***d*** *EEG activity in recognized (green line) and in unrecognized trials (orange line) locked with target onset (low contrast study). The gray horizontal line represents the significant time windows where EEG activity is significantly different for the two conditions (p < 0.05).* ***e*** *Scalp topography – average CNV amplitude in correct and in incorrect trials, respectively (backward masking study).* ***f*** *Scalp topography – average CNV amplitude in recognized and in unrecognized trials, respectively (low contrast study).* ***g*** *Scalp topography – differences of the CNV amplitude between correct and incorrect trials (backward masking study).* ***h*** *Scalp topography – differences of the CNV amplitude between recognized and unrecognized trials (low contrast study). In graphs* ***a****,* ***b****,* ***c*** *and* ***d****, data are represented as mean ± standard error of the mean across participants.*

3.2.6.2 CNN: classification of visual performance based on pre-stimulus neural activity

After studying how EEG pre-stimulus activity was related to task performance in both studies, we then investigated if the classifier was able to use EEG pre-stimulus features to predict the participant's performance. Considering that we used the EEG signal, which is characterized by being a continuous signal, we realized that we needed a more powerful classifier capable of receiving a relatively high amount of data as input. That said, we decided to use the EEGNet as it can extract spatial and temporal features of the EEG for the classification. Then, we incorporated the parameters of each physiological signal into the classifier. In this case, we used the multimodal classifier.

In both Figure 3.23 – a) and Figure 3.23 – b), it is evident that, with the pre-stimulus EEG, we were able to predict the participant's performance, in both studies. Despite the fact that no significant differences were observed between the neural activity and trial types (correct/incorrect trials in the backward masking study and recognized/unrecognized trials in the low contrast study), the classifier was able to use this information to correctly classify the trials. These findings align with the results reported in previous studies (10), suggesting an association between EEG activity and the recognition of visual stimuli.

We then incorporated the different physiological signals into the classifier and assessed how these influenced its predictive capacity. If the physiological signal had a direct influence on visual processing, then their pre-stimulus state might contribute towards predicting the participants visual performance. In both Figure 3.23 – a) and Figure 3.23 – b), we can see that the incorporation of different signals produced different behaviors in both studies. In the backward masking study, the incorporation of the various physiological signals resulted in lower performances compared to when using only the pre-stimulus neural activity. In fact, incorporation of both saccadic activity and the combination of all physiological signals led to statistically inferior performance. On the contrary, in the low contrast study, we can see that the incorporation of physiological signals yielded average performances that were not significantly different from the performance of the classifier using only the pre-stimulus EEG activity.

	Backward masking study	Low contrast study
Only EEG pre-stimulus	**80%**	**80%**
EEG + ECG	**66.67%**	**100%**
EEG + Respiration	**73.33%**	**100%**
EEG + Pupil	**66.67%**	**100%**
EEG + Blinks	**93.33%**	**100%**
EEG + Saccades	**66.67%**	**100%**
EEG + All measures	**60%**	**100%**

Table 18 *— Percentage of classifiers using the physiological data and neural activity as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	Low contrast study
EEG + ECG	$t\,(14) = 2.010$; $p = 0.064$	$t\,(14) = -0.359$; $p = 0.725$
EEG + Respiration	$t\,(14) = 1.627$; $p = 0.126$	$t\,(14) = -0.132$; $p = 0.897$
EEG + Pupil	$t\,(14) = 1.442$; $p = 0.171$	$t\,(14) = -0.696$; $p = 0.498$
EEG + Blinks	$t\,(14) = 1.449$; $p = 0.169$	$t\,(14) = -1.817$; $p = 0.091$
EEG + Saccades	$\mathbf{t\,(14) = 2.475; p = 0.027}$	$t\,(14) = 0.058$; $p = 0.954$
EEG + All measures	$\mathbf{t\,(14) = 3.130; p = 0.007}$	$t\,(14) = -0.914$; $p = 0.376$

Table 19 *– Comparison between the performance achieved using only neural activity and the performance attained through the combination of neural activity with different physiological signals. We used the paired-sample t-test to evaluate whether each physiological signal contains supplementary information that can be extracted by the classifier. The statistically significant analyses are highlighted.*

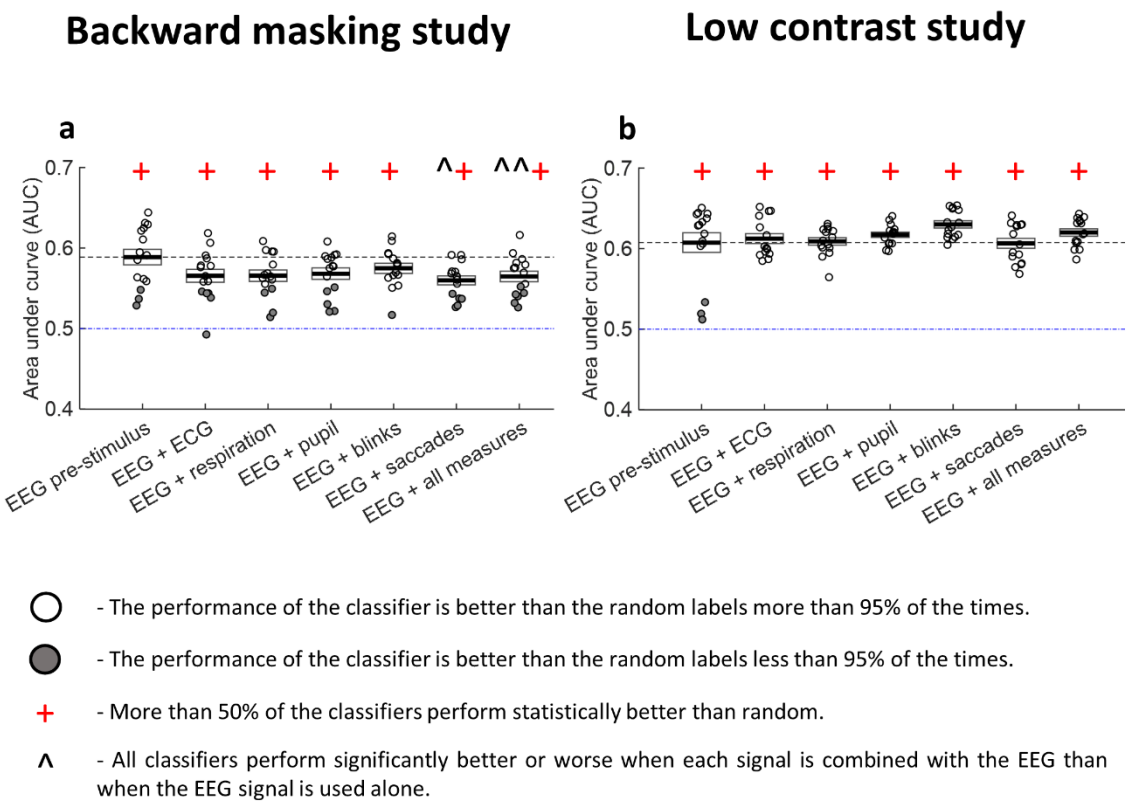


Figure 3.23 *– The neural pre-stimulus activity influences stimulus detection. Integrating pre-stimulus body activity did not significantly improve the classifier's performance.* ***a*** *AUC using a CNN model to predict trial performance (backward masking study).* ***b*** *AUC using a CNN model to predict stimulus detection (low contrast study). The 1st bar shows the results for when the neural activity pre-stimulus is used. The remaining bars concern the performance of the classifiers combining each physiological signal pre-stimulus with neural activity. The black line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this pre-stimulus activity as input, it results in a performance that is statistically different to the performance produced when only neural pre-stimulus activity is used. ^p<0.05, ^^p≤0.01, ^^^p≤0.001.*

These results suggest that the physiological signals do not convey additional information that is not present already in the pre-stimulus EEG signal and therefore do not contribute to a better performance of the classifier. However, it is possible that this particular multimodal classifier is not as efficient as the SVM at using the physiological features to classify the participants visual performance. In order to ascertain the capacity of the multimodal classifier to extract activity patterns from physiological signals, we implemented a control algorithm that uses the multimodal classifier while setting the EEG to zero. It should be noted that, in this analysis, we focused only on the combination of all parameters of each physiological signal, rather than studying each parameter individually.

In the backward masking study, none of the physiological signals contained activity patterns that allowed the multimodal classifier to predict visual performance (Figure 3.24 –

a)). As demonstrated in Table 20, only a maximum of 26.67% of the classifiers showed behavior superior to random. Conversely, in the low contrast study, it was possible to predict visual performance through pre-stimulus body activity. Unlike what was seen with the SVM algorithm, where the classifier was able to predict visual performance using pupillary response and respiratory activity, with the multimodal classifier, only the pupillary response allowed the classifier to extract activity patterns (Figure 3.24 – b)). Moreover, the combination of all physiological signals also yielded the ability to predict visual performance. These results suggest that only the pupillary response provided information to the classifier that allowed visual performance to be predicted. Nonetheless, as depicted in Figure 3.23– b), pupil activity did not provide extra information to the classifier above the information that was already present in the EEG. One of the hypotheses that can justify this result could be the fact that pupil activity is somehow reflected in the EEG signal (the correlation matrix between each pupil measure and the EEG signal can be found in the appendix). These results make sense, as it is anticipated that the brain controls and monitors the body's state. For the remaining physiological signals, we were unable to draw conclusions. This is due to the failure of the classifier to extract activity patterns from these signals, making it impossible to determine whether or not these signals contain extra information that is not reflected in the EEG signal.

	Backward masking study	**Low contrast study**
ECG	0%	13.33%
Respiration	0%	0%
Pupil	0%	**86.67%**
Blinks	26.67%	0%
Saccades	6.67%	20%
All measures	20%	**53.33%**

Table 20 *— Percentage of classifiers using only physiological data as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	Low contrast study
ECG	***t* (14) 5.128; *p* <0.001**	***t* (14) = 3.384; *p* = 0.004**
Respiration	***t* (14) = 5.754; *p* <0.001**	***t* (14) = 6.1908; *p* <0.001**
Pupil	***t* (14) = 3.6181; *p* = 0.003**	*t* (14) = -1.729; *p* = 0.106
Blinks	*t* (14) = -1.378; *p* = 0.190	***t* (14) = 9.585; *p* <0.001**
Saccades	*t* (14) = 0.709; *p* = 0.490	***t* (14) = 3.672; *p* = 0.002**

Table 21 *– Comparison between the performance achieved using each physiological signal individually and the performance attained through the combination of all physiological signals. We used the paired-sample t-test to evaluate the contribution of each signal when all physiological signals were combined. The statistically significant analyses are highlighted.*

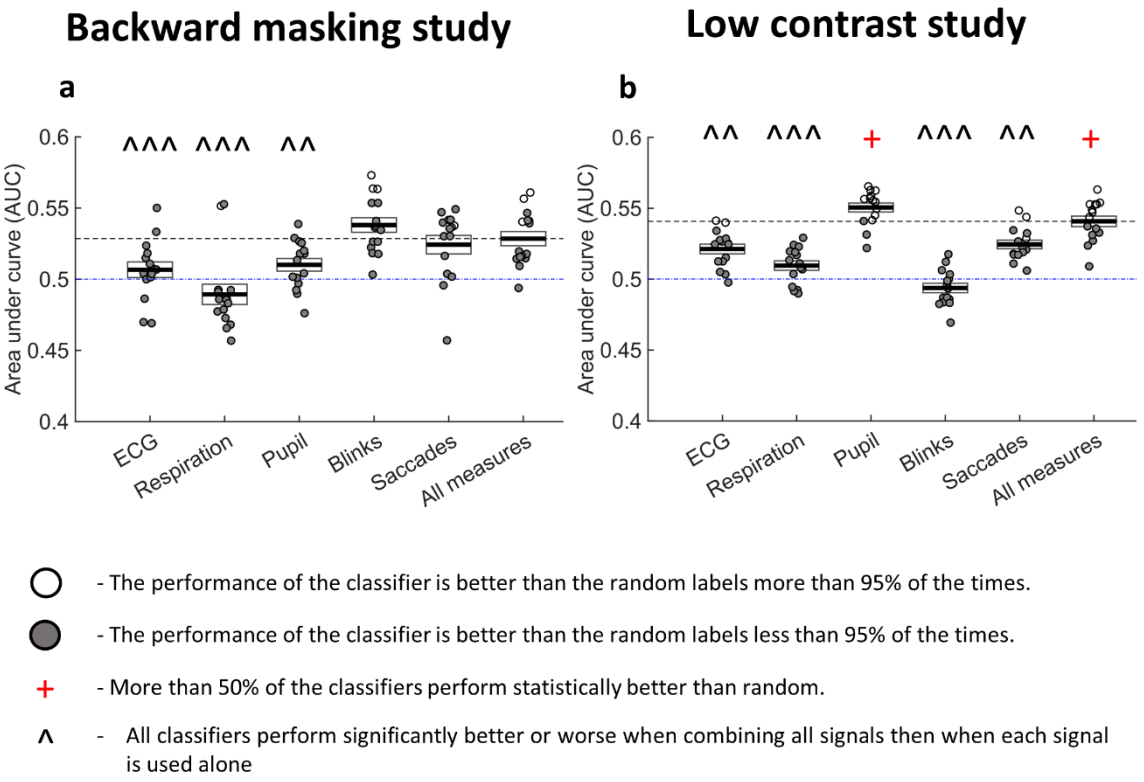


Figure 3.24 *– Multimodal classifiers' performance using only physiological pre-stimulus activity. In the backward masking study, none of the signals allowed us to predict trial performance. Conversely, in the low contrast study, through pupil activity and combining all physiological signals we predicted stimulus detection.* ***a*** *AUC using a CNN model to predict trial performance (backward masking study).* ***b*** *AUC using a CNN model to predict stimulus detection (low contrast study). The last bar shows the results for when all physiological signals are used. The remaining bars concern the performance of the classifiers using each physiological signal individually. The black line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this pre-stimulus activity as input, it results in a performance that is statistically different to the performance produced when all physiological signals are used. ^p<0.05, ^^p≤0.01, ^^^p≤0.001.*

Comparing both the low contrast study and the backward masking study, it becomes apparent that pre-stimulus neural activity influences visual processing. Furthermore, there appears to be an association between pre-stimulus body activity and visual performance. However, our analyses did not allow us to conclude whether this pre-stimulus body activity has a direct impact on sensory processing or whether it is only associated with brain state fluctuations that control perception.

3.3 Study of the influence of pre-stimulus brain and body activity on visual stimulus neural representations

After checking the ability of the classifier to predict visual performance from pre-stimulus physiological signals, we decided to study its ability to discriminate the visual stimulus category from the post-stimulus EEG evoked responses and how these responses might be influenced by pre-stimulus brain and body state.

We found that, in the backward masking study, the visual stimuli evoked ERPs that were significantly different between cars and houses (Figure 3.25 – a)). In Figure 3.25 – a), we can see that these differences occur particularly around 170 ms after stimulus onset, where the negative waveform that peaked approximately 170 ms after stimulus presentation (N170) was more negative for car stimuli than for house. This observation is in agreement with the conclusions verified in (130). In contrast, the low contrast study reveals no significant differences in the ERPs evoked by the car and house stimuli. In Figure 3.25 – b), we can see that the negative waveform achieved its maximum amplitude approximately 200 ms after stimulus onset, rather than the observed 170 ms after stimulus onset in the backward masking study. The observed delay in stimulus processing might be attributed to the reduction in stimulus contrast.

In the backward masking study, we observe significant differences in the EEG signal between car and house trials within the time interval associated with neural processing. However, this pattern is not verified in the low contrast study. In this study, we can see very similar patterns of activity in the conditions examined. This difference may be related to changes in task difficulty, resulting in changes in neural processing. Allied to this, we also believe that the decrease in contrast hinders the visualization of certain characteristics, which could be responsible for the generation of distinct EEG signals. It should be noted that, in this thesis, we present the signal relative to the PO7 channel as a representation. However, none of the remaining channels exhibited significant differences in processing the two trial

types. Some channels, such as the AF4 channel, did exhibit temporal points wherein the signal differed between the two types of trials, but these points were situated around the 400ms mark, which might be associated with decision-making processes (Figure 3.25 – c) and Figure 3.25 – d)).

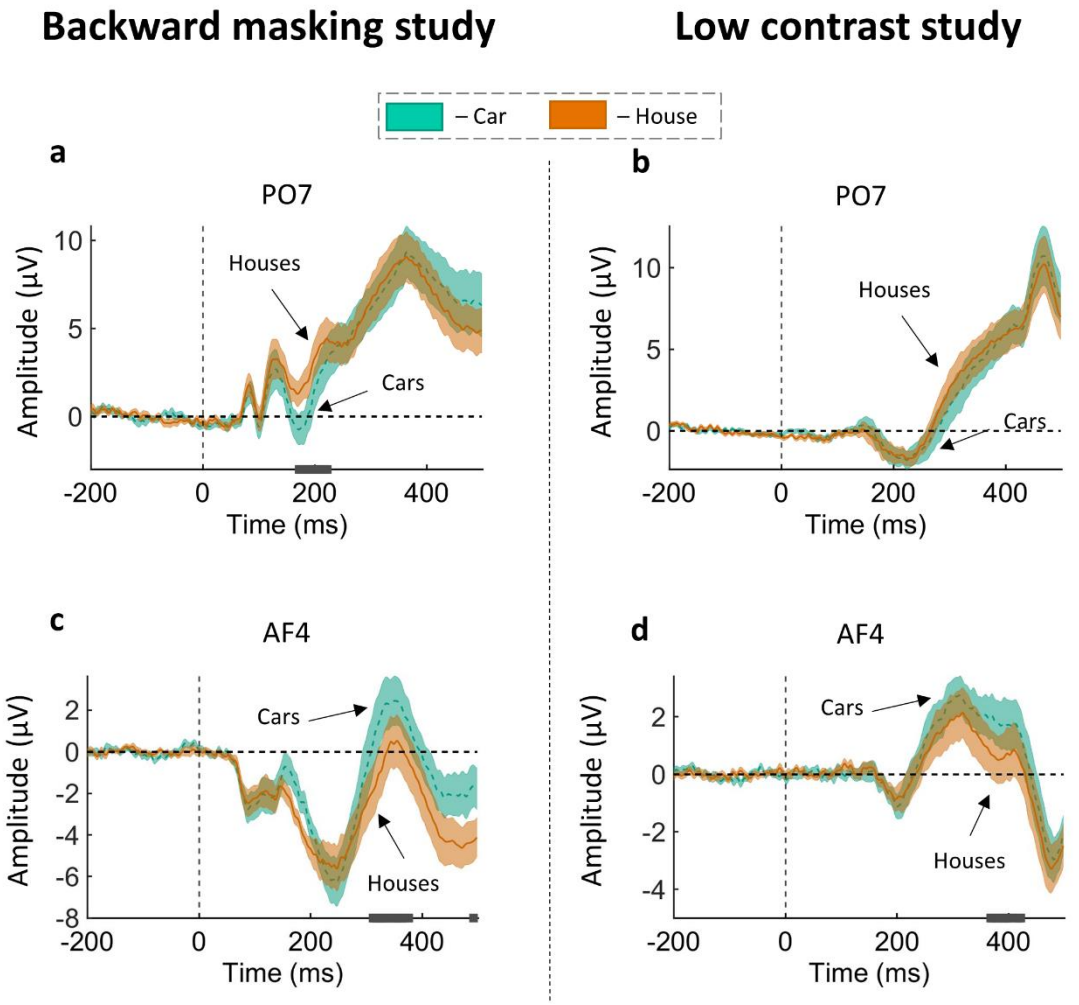


Figure 3.25 *– The visual stimuli induced significant differences between car and house trials. In the backward masking study, we can see significant differences within the time interval associated with neural processing and decision-making processes. In the low contrast study, in turn, we only can see significant differences within the time interval associated with decision-making processes.* ***a and b*** *Target-locked EEG signal recorded from electrode PO7 in car and house trials (a – backward masking study; b – low contrast study).* ***c and d*** *Target-locked EEG signal recorded from electrode AF4 in car and house trials (a – backward masking study; b – low contrast study). The gray horizontal line represents the time windows where the EEG signal is significantly different for the two conditions ($p < 0.05$). In all graphs, data are represented as mean ± standard error of the mean across participants.*

Next, we investigated if the classifier was able to recognize the patterns associated with the different categories and, in this way, correctly classify stimulus category. Thus, the classifier was fed with the ERP in each trial as well as the respective trial category.

It is important to highlight that the metric selected for the analyses conducted throughout the thesis is the AUC. However, as in this analysis the classes are balanced, we are going to compare the participant's performance with the classifier's performance, we decided to choose accuracy as the metric for a more precise comparison.

Similar to the previous analyses, we are going to analyze the backward masking study and then the low contrast study. Analyzing the Figure 3.26 – a), we can see that the accuracy of the classifier for each participant is significantly above 50% ($t(15)$ = -7.296, p <0.001). We hypothesized that a higher participant's accuracy was correlated with a higher accuracy of the classifier. This hypothesis is based on the fact that if the participant has a higher accuracy means that he saw more trials, subsequently resulting in more pronounced ERPs. In this way, we performed a correlation between the accuracy of the participants and the accuracy of the classifier. The correlation was not statistically significant (r = 0.186; p = 0.489). In the low contrast study, on the contrary, in three participants, the accuracy of the classifier is less than 50% (Figure 3.26 – b)), nevertheless the average across participants was significantly higher than 50% ($t(18)$ = -3.726, p = 0.002). Interestingly, the correlation between participant's accuracy and classifier's accuracy was statistically significant (r = 0.488; p = 0.034). Comparing the two studies, as anticipated, the classifier presents a superior performance for the backward masking study, due to the generation of more pronounced ERP patterns in this particular study.

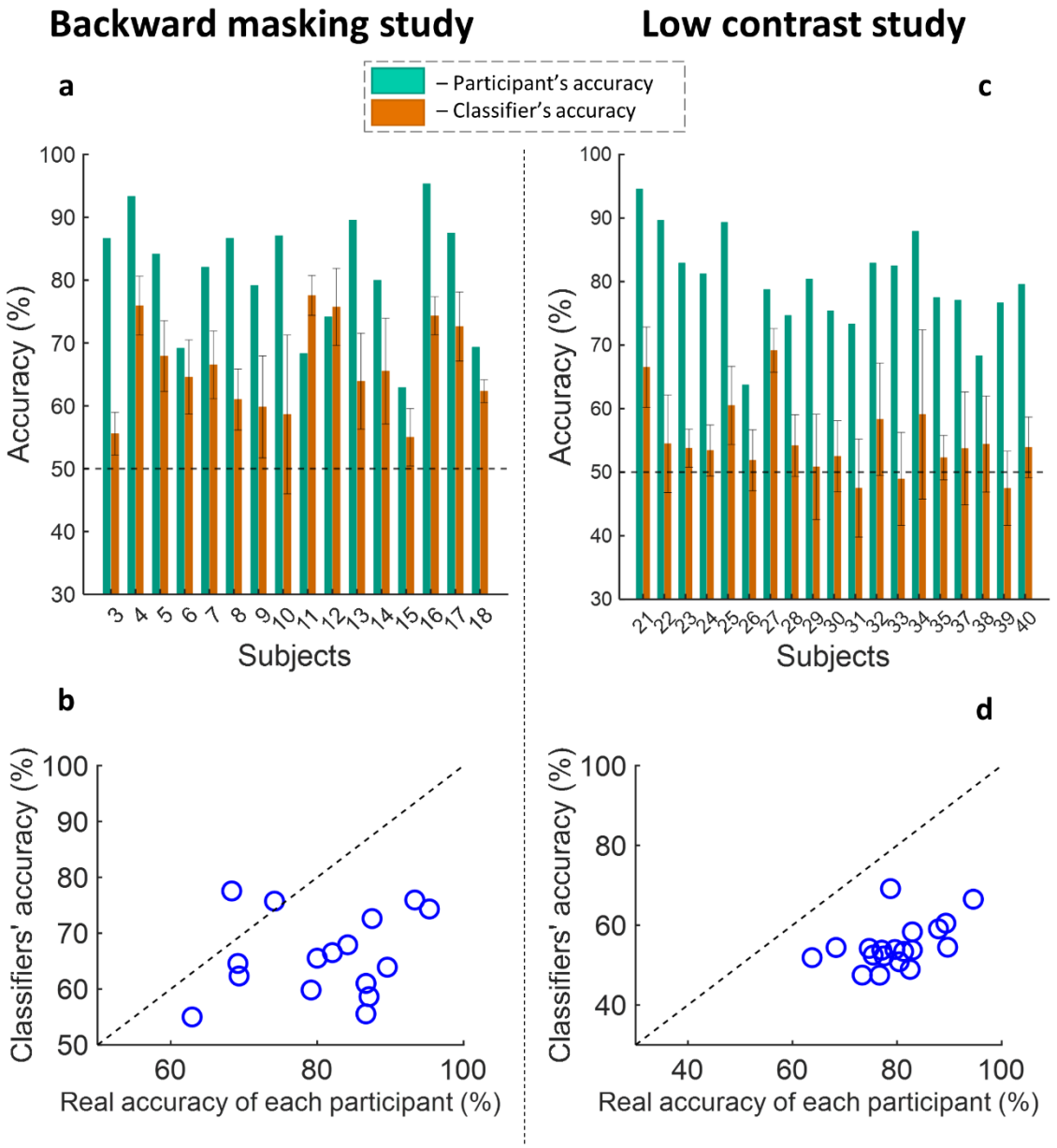


Figure 3.26 *– Post-stimulus EEG allows predicting the trial category.* ***a*** *Comparison between the EEGNet accuracy and participant's accuracy (backward masking study).* ***b*** *No significant correlation between classifier's accuracy and participant's accuracy (backward masking study).* ***c*** *Comparison between performance of the EEGNet accuracy and participant's accuracy (low contrast study).* ***d*** *Correlation between classifier's accuracy and participant's accuracy (low contrast study). The green bars concern the participant's accuracy and the orange bars concern the classifier's accuracy. The error bar represents the standard error of the mean. In sub-figures* ***b*** *and* ***d****, each circle represents a participant.*

Finally, we proceeded to investigate whether pre-stimulus neural and physiological activity modulates stimulus decoding. To this end, we incorporated body and brain pre-stimulus activity into the classifier. We included this pre-stimulus activity to investigate whether this activity modulates the neural representation of the stimulus. If this modulation occurs, the classifier can learn this relation. As an illustration, a given pre-stimulus

fluctuation could lead to the modulation of stimulus representation, inducing a more pronounced waveform, thereby facilitating the classifier's learning process.

In both Figure 3.27 – a) and Figure 3.27 – b), it is evident that with the ERPs we were able to extract activity patterns associated with neural processing and decision-making processes, thereby facilitating the discrimination of stimulus categories. We can see that, in both studies, more than 50% of the classifiers exhibited a performance statistically superior to random, confirming the efficacy of the classifier to discriminate stimulus category (Table 22). In the same figures, we can observe that the incorporation of the pre-stimulus physiological and neural activity did not produce an improvement in the classifier's performance. In fact, the combination of certain physiological signals with ERPs results in statistically lower performances compared to the ones observed when the post-stimulus activity was used as input. These results suggest that the pre-stimulus activity does not modulate the neural representation of the stimulus.

Our project was motivated by the results presented in (1) and (10). In (1), the authors verified that the pre-stimulus neural activity modulates the neural representation of the stimulus, thereby enhancing the classifier's learning process. In (10), in turn, the authors detected the presence of fluctuations that facilitate the recognition of one stimulus over another. In summary, both studies established the presence of pre-stimulus neural activity fluctuations that influence the processing according to their respective categories. However, as illustrated in Figure 3.27, the incorporation of the pre-stimulus neural activity resulted in a performance that is statistically inferior to the performance achieved by the classifier when exclusively using ERPs. These observed discrepancies between our study and the findings presented in (1) and (10) can be attributed to various factors:

- EEG has a lower spatial resolution than MEG or intracranial recordings, making it difficult to distinguish fluctuations in neural activity in different brain regions.

- the feature selected for integration into the classifier might not have been optimal. To incorporate in the classifier, we opted to incorporate only the average activity of each channel during the second preceding stimulus presentation. However, this approach might not have been the most appropriate, as the process of averaging could have masked specific activity fluctuations, potentially resulting in a biased behavior.

- it is plausible that the task selected for this study might not have been the best, potentially hiding the impact of pre-stimulus activity fluctuations.

In conclusion, contrary to what was verified in (1) and (10), the integration of pre-stimulus neural activity did not significantly modulate the decoding process of the stimulus. From these findings, we can conclude that pre-stimulus activity influences stimulus recognition, but it did not appear to influence the way that the stimulus is decoded.

	Backward masking study	**Low contrast study**
ERP	**100%**	**93.33%**
ERP + EEG pre-stimulus	**100%**	**86.67%**
ERP + ECG	**100%**	**86.67%**
ERP + Respiration	**100%**	**86.67%**
ERP + Pupil	**100%**	**100%**
ERP + Blinks	**100%**	**93.33%**
ERP + Saccades	**100%**	**100%**
ERP + All measures	**100%**	**93.33%**

Table 22 *- Percentage of classifiers using the combination of body and brain pre-stimulus activity with ERPs that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance. The successful classifications are highlighted.*

	Backward masking study	**Low contrast study**
ERP + EEG pre-stimulus	***t (14)* = 3.352; *p* = 0.005**	***t (14)* = 2.649; *p* = 0.019**
ERP + ECG	*t (14)* = 0.873; *p* = 0.397	*t (14)* = 1.151; *p* = 0.269
ERP + Respiration	*t (14)* = 1.271; *p* = 0.224	*t (14)* = 0.744; *p* = 0.469
ERP + Pupil	*t (14)* = -0.403; *p* = 0.693	*t (14)* = -0.405; *p* = 0.692
ERP + Blinks	*t (14)* = 1.578; *p* = 0.137	*t (14)* = 1.039; *p* = 0.316
ERP + Saccades	*t (14)* = 1.957; *p* = 0.071	*t (14)* = -0.865; *p* = 0.402
ERP + All measures	*t (14)* = 1.800; *p* = 0.094	*t (14)* = 1.866; *p* = 0.083

Table 23 *- Comparison between the performance achieved using only ERPs and the performance attained through the combination of ERPs with different pre-stimulus physiological and neural activity. We used the paired-sample t-test to evaluate whether each pre-stimulus activity modulates how stimulus is decoded. The statistically significant analyses are highlighted.*

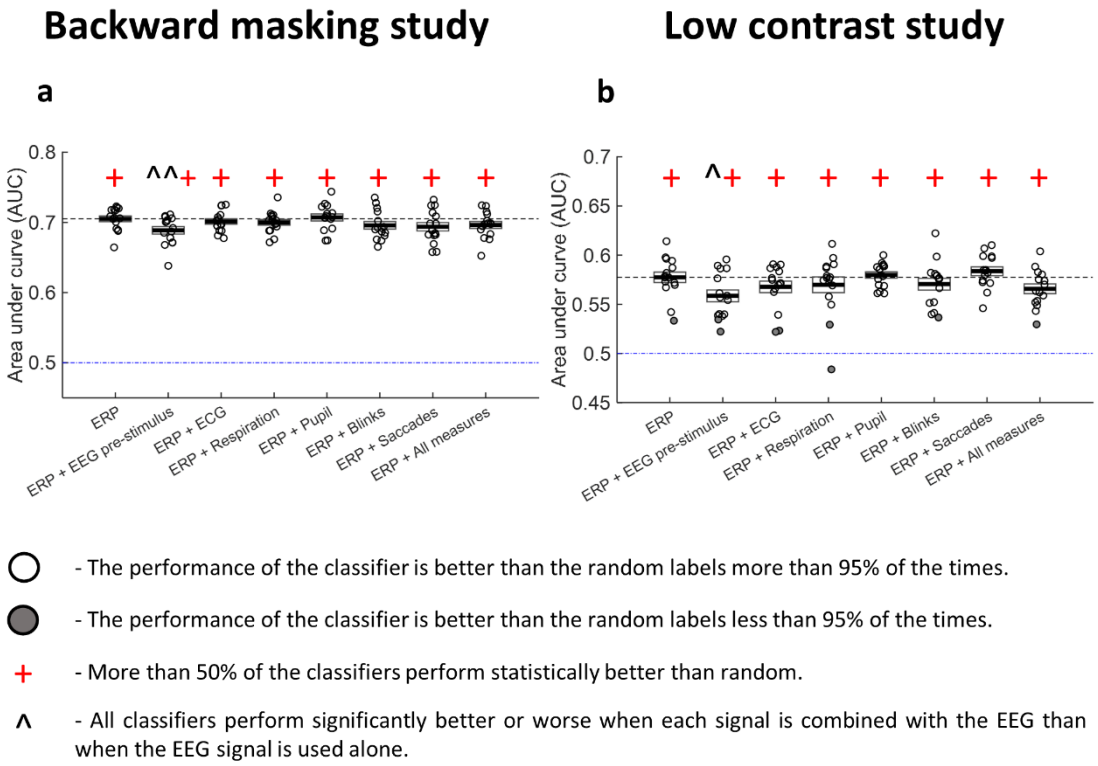


Figure 3.27 *– There is no evidences that the brain and body physiological pre-stimulus activities modulate visual stimuli neural representations.* ***a*** *AUC using a CNN model to discriminate stimulus category (backward masking study).* ***b*** *AUC using a CNN model to discriminate stimulus category (low contrast study). The 1st bar shows the results for when only ERPs are used. The remaining bars concern the performance of the classifiers combining pre-stimulus physiological and neural activity with the ERPs. The black line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this pre-stimulus activity as input, it results in a performance that is statistically different to the performance produced when only ERPs are used. ^p<0.05, ^^p≤0.01, ^^^p≤0.001*

3.4 Limitations of our project

As observed in the previous section, the SVM algorithm presented a greater ability to predict visual performance using only physiological features than the multimodal classifier. When defining the multimodal algorithm, most of the trainable parameters were found in the blocks responsible for processing the EEG part, while the remaining parameters were adjusted based on the physiological signals. As a result, the algorithm gave too much weight to the EEG part, resulting in almost the totality of the weights trained on zero-valued images (in the control analysis). Having said that, we believe that the way that this multimodal algorithm was developed was probably not the best and future work should explore different architectures. Allied to this, the SVM algorithm is probably more robust to the type of inputs employed in the study of physiological pre-stimulus activity. In the future, one way to use the

multimodal classifier only with physiological activity, without containing useless information that the EEGNet extracted from the EEG (even if set to zeros), is to use a simpler classifier that is composed only of the layers of the multimodal classifier that process physiological activity.

This limitation might algo affect the multimodal classifier when combining continuous EEG signal with the physiological features. As previously stated, most of the trainable parameters are in the layer blocks responsible for processing the EEG, while a small minority are in the layer that process physiological activity. In this way, the multimodal classifier ends up weighing more the information extracted from the EEG than the information extracted from the body signals. This behavior can be one of the hypotheses for why adding physiological signals does not significantly improve the classifier's performance. One of the ways to combat this problem is to, together with the continuous EEG signal, perform time-series analyses on physiological activity. With this analysis, we believe that the classifier can assign equal importance to neural and physiological activity.

4 Conclusions

4.1 General conclusions

In this project, we aimed to investigate the influence of pre-stimulus activity, both neural and physiological, on visual perception and visual processing. To achieve this, we used two different versions of a visual discrimination task, the backward masking task and the low contrast task. In the backward masking task, the only information we had was the accuracy in a given trial, that is, whether or not the person had responded correctly according to the stimulus category. However, these types of responses have some limitations. For example, on some occasions, it was possible that the participants guessed the category correctly purely by chance, without recognizing the stimulus. Consequently, the neural and physiological states during these trials might not have favored stimulus recognition yet the analyses were done considering those trials as the trials where recognition had occurred. That said, we consider that the backward masking study was too noisy, whereas the low contrast study provided more informative and relevant results due to enhanced design and additional features. As a result, we decided to focus our conclusions on the findings derived from the low contrast study.

With the changes in the design of the task, our intention was to create a difficult task, where the trials would be at the threshold of visual recognition. In fact, we aimed for participants to recognize approximately 50% of the trials. This would enable us to observe a clear influence of the state of expectation on stimulus recognition. However, overall, participants recognized slightly more trials than expected. This result may come from the fact that the conditions of the adaptive procedure were different from the conditions observed during the actual task. Allied to this, some participants may not be as attentive during the adaptive procedure as they were during task performance, which may have influenced the contrast level corresponding to a subjective recognition of 50%.

Following the behavior analyses, it was time to assess how the pre-stimulus activity influenced stimulus recognition. To do so, we started by analyzing the influence of cardiac activity. We started by verifying that an auditory cue induced cardiac deceleration. However, although there were significant differences in heart rate at the time of stimulus presentation, these fluctuations in heart rate were not significantly associated with the participant's recognition of the stimulus. In fact, the classifier could not accurately predict stimulus recognition based on cardiac activity. These results turned out to be contradictory and suggest that differences between the classifier's algorithms and traditional statistical analyses might lead to different conclusions and care should be taken when focusing on only one type of analyses.

Next, we went to study how breathing activity was related to visual processing. Initially, we observed that the auditory cue was followed by an increase in the duration of the respiratory cycle. Allied to this, we verified that the cycle with the maximum duration was the cycle containing the stimulus presentation. We believe this is a novel result that has not been observed previously. Moreover, pre-stimulus respiratory activity had information that could predict visual performance. In fact, despite the fact that no significant differences were observed between the respiratory activity in recognized and unrecognized trials, the classifier was able to extract patterns of activity that enabled the prediction of stimulus recognition. These results suggest that the simple analysis of the average, as we have seen for cardiac activity, can be biased, as it does not take into account the behavior along the trials. Thus, the results suggest that respiratory activity somehow is associated with visual stimulus recognition.

As anticipated, the alerting stimulus induced pupil dilation. This is an expected behavior since pupil dilation reflects the arousal levels. As verified in other studies, pupil size and behavior exhibited significant differences between trials where stimulus was recognized and those where it was unrecognized. Moreover, when incorporating the pupillary response data into the classifier, we were able to successfully predict stimulus detection. These results suggest that pupil-linked arousal is significantly associated with visual perception.

Finally, we investigated the influence of blinking and saccadic activity on stimulus detection. In the state of expectation there is a significant decrease in the occurrence of both blinks and saccades, minimizing loss of visual information. In the analyses of blinks and saccades, we observed that when they occurred very close to the stimulus, participants tended not to recognize the stimulus; however, this was significant in the saccades but not in the blinks analyses. Despite having this influence, the occurrence of blinks and saccades near to the stimulus is relatively rare due to the reduced occurrence during the state of expectation. Allied to this, there were specific activity patterns only when a blink or a saccade occurred very close to the stimulus presentation, whereas no detectable activity patterns were evident in the remaining trials. This could have affected the ability of the classifier to learn from these data. In fact, it was found that, both with the activity of blinks and saccades, the classifier was not able to predict stimulus recognition. In conclusion, our findings demonstrate that while blinks and saccades might influence stimulus detection when occurring close to the stimulus, this influence is sporadic.

In addition, we also investigated pre-stimulus neural activity. The findings suggested that there are no significant differences in EEG cortical activity between trials where the participant recognizes the stimulus and those in which he does not. These results were contradicted by the classifier's performance when using only the pre-stimulus EEG activity as input, where we can see that it is possible to extract excitability patterns from the cortex,

enabling the prediction of stimulus detection. Consequently, these findings suggest that neural activity is associated with visual stimulus recognition.

Our primary objective was to examine how the pre-stimulus activity modulates cortical processing of visual stimuli and subsequently affects visual perception. In addition to analyzing the pre-stimulus neural activity, we also investigated the influence of the physiological state before the stimulus presentation. In order to assess the multimodal classifier's ability to extract patterns from different physiological signals, it was necessary to set the EEG to zero, thereby ensuring that the classifier's predictions were only influenced by pre-stimulus body activity. The multimodal classifier was only able to predict visual performance when using pupillary activity as input and when incorporating all the physiological signals. Applying statistical tests, we found that, probably, when incorporating all the physiological signals, it was the pupillary activity that provided the activity patterns to the classifier. Summarizing, the multimodal classifier was only able to extract patterns from the pupil.

After evaluating this control algorithm, we integrated various measures with neural activity with the expectation of extracting extra information that the classifier cannot extract from neural activity. This analysis aimed at understanding if the physiological signals might affect visual perception independently from the pre-stimulus activity already observed in the brain. The combination of several physiological signals with neural pre-stimulus activity did not result in significant improvements in the classifier's performance. Considering that the control algorithm was only able to extract patterns of pupil activity, it was anticipated that only this signal would bring supplementary information to the classifier. However, pupil activity did not provide extra information. This result suggests that pupil activity is somehow already reflected in the EEG signal. We believe that this result may be related to the fact that our brain controls and is aware of our body, which causes a correlation between neural and physiological activity. For the remaining physiological signals, we were unable to draw conclusions. One of the hypotheses that may also have limited the results was the choice of the multimodal algorithm, which gives much more weight to the EEG signal than to the features of the remaining physiological signals.

Upon confirming that pre-stimulus activity influences the recognition of the stimulus regardless of the category, we went to verify whether this influence was specific to different stimulus categories. That is, we went to check if the fluctuations in physiological and neural state modulated the encoding of each visual category in the brain. To achieve this, we combined the ERPs induced by the visual stimulus with pre-stimulus physiological and neural activity in a single multimodal classifier. However, the inclusion of this pre-stimulus activity did not produce significant improvements in the classifier's performance. This outcome suggests that the pre-stimulus activity does not modulate the decoding process of

the stimulus. This contradicts what was found in (1) and (10). We believe that the discrepancies on the results might be attributed to the fact that EEG has a lower spatial resolution than MEG or intracranial recordings. Allied to this, the selected features to be included on the classifier and the employed task might not have been optimal. Finally, it is plausible that the chosen algorithm, as mentioned previously, also has an influence on these results.

Throughout this project, we conducted a comprehensive investigation into the influence of pre-stimulus activity, both neural and physiological aspects, on visual stimulus detection. The results demonstrate that pre-stimulus activity indeed plays a crucial role in the detection process. However, this influence does not appear to modulate stimulus representation in the brain. In other words, pre-stimulus activity exerts a general influence on the stimulus recognition, but it has no differential effects on recognition of different stimuli. Moreover, we did not find any evidence that body physiological signals affect visual perception directly.

4.2 Future work

One of the constraints of our project concerns the inclusion of single measures of the different physiological signals (at a fixed time point), rather than incorporating the signal's actual time course. Consequently, when extracting the measures, certain fluctuations in the signal may be "ignored", which may influence the neural representations of the stimuli as observed in previous studies (1,10). Therefore, in the future, it would be important to incorporate the time course of the different physiological signals preserving the signals time dynamics and not just a single measure of the signal.

We observed that visual perception was associated with fluctuations in body physiology. However, we did not study how body physiology is associated with visual processing in the cortex. It would be interesting to extend these findings to analyze the relationship between body physiology and the ERPs. Moreover, through EEG source analyses it would be interesting to focus our study in the specific areas of the visual cortex to further understand the mechanisms involved. Furthermore, there is still much to be explored regarding the interplay between the body and the brain and their mutual influence on sensory processing.

5 Appendix

5.1 Pre-stimulus neural and physiological activity

Backward masking study

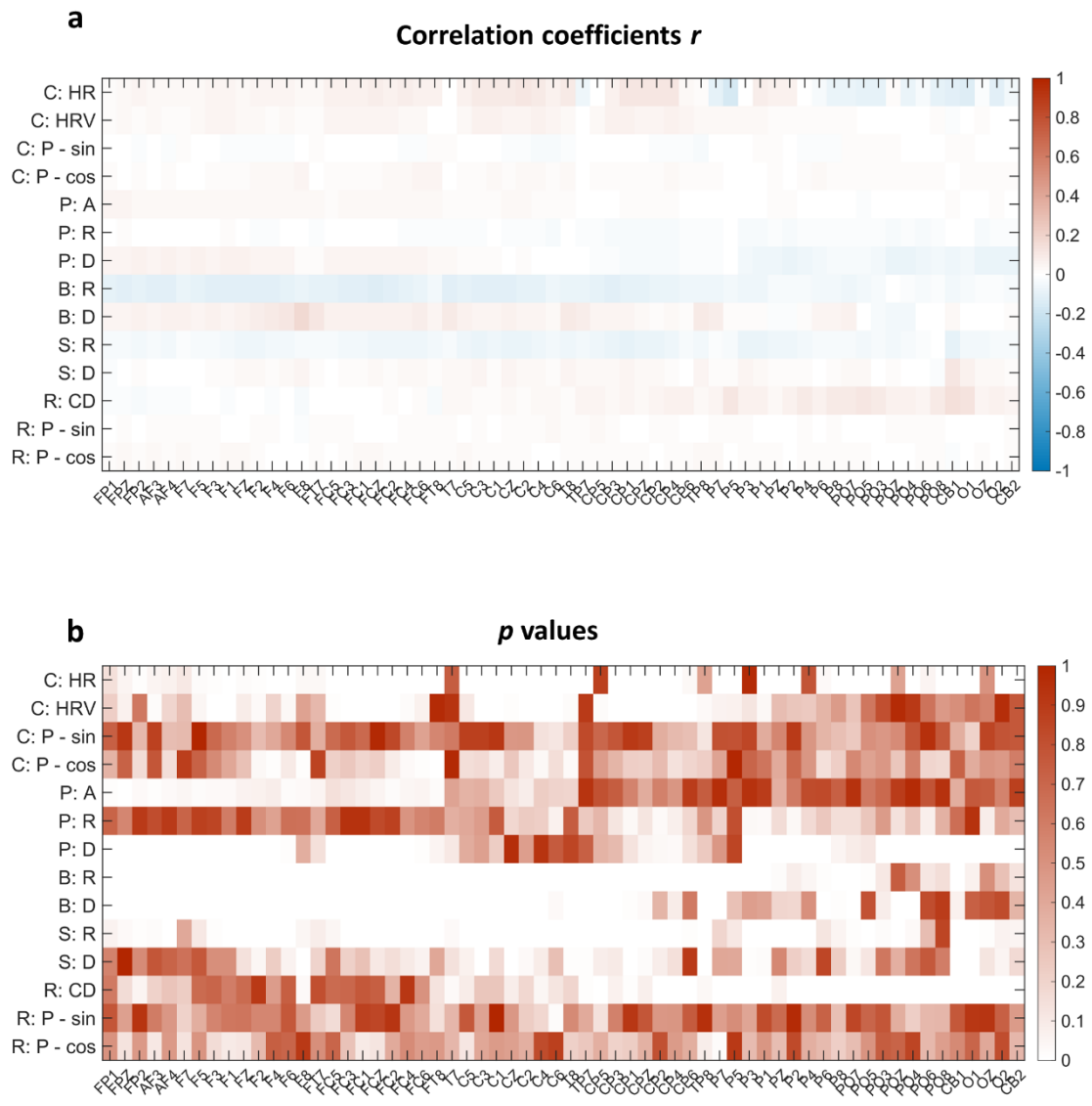


Figure 5.1 *– Correlation between the different measures and EEG activity from each channel. Several studies suggest that our brain is aware of the physiological state of our body. That said, it is to be expected that the behavior of each measure is reflected in the EEG signal. Thus, in this analysis, we have the correlation between each measure and the electrical activity of each electrode.* ***a*** *Correlation coefficients.* ***b*** *Significance values (p-values) of the correlation. It should be noted that this figure represents only the correlation and significance values for backward masking study. C: HR – Heart rate; C: HRV – Heart rate variation; C: P – sin – sine component of the cardiac phase; C: P – cos – cosine component of the cardiac phase; P: A – pupil size; P: R – relative pupil size; P: D – average pupil derivative; B: R – blink rate; B: D – distance of the last blink; S: R – saccades rate; S: D – distance of the last saccade; R: CD – respiratory cycle duration; R: P – sin – sine component of the respiratory phase; R: P – cos – cosine component of the respiratory phase.*

Low contrast study

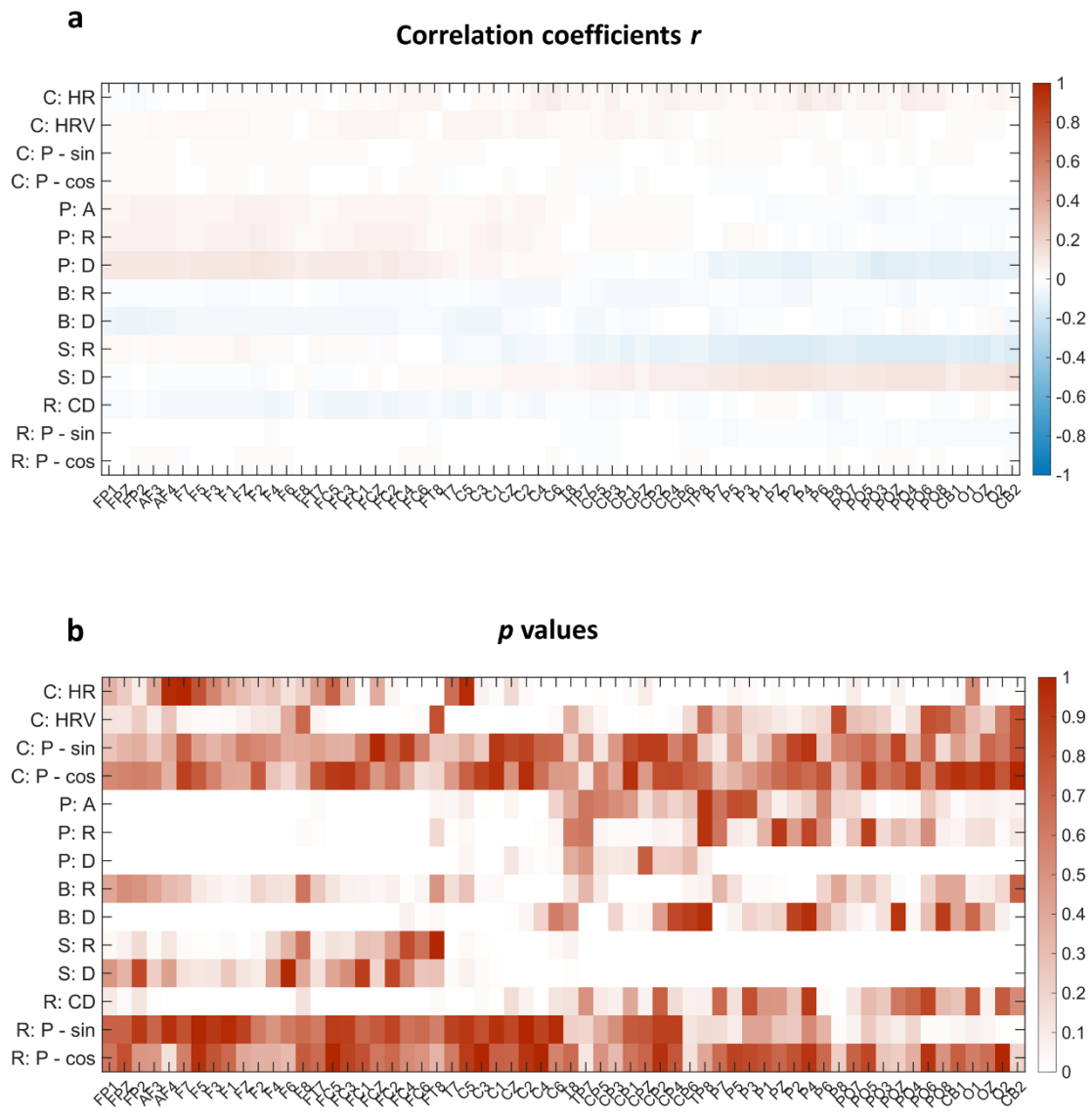


Figure 5.2 *- Correlation between the different measures and EEG activity from each channel. Several studies suggest that our brain is aware of the physiological state of our body. That said, it is to be expected that the behavior of each measure is reflected in the EEG signal. Thus, in this analysis, we have the correlation between each measure and the electrical activity of each electrode.* ***a*** *Correlation coefficients.* ***b*** *Significance values (p-values) of the correlation. It should be noted that this figure represents only the correlation and significance values for the low contrast study. C: HR – Heart rate; C: HRV – Heart rate variation; C: P – sin – sine component of the cardiac phase; C: P – cos – cosine component of the cardiac phase; P: A – pupil size; P: R – relative pupil size; P: D – average pupil derivative; B: R – blink rate; B: D – distance of the last blink; S: R – saccades rate; S: D – distance of the last saccade; R: CD – respiratory cycle duration; R: P – sin – sine component of the respiratory phase; R: P – cos – cosine component of the respiratory phase.*

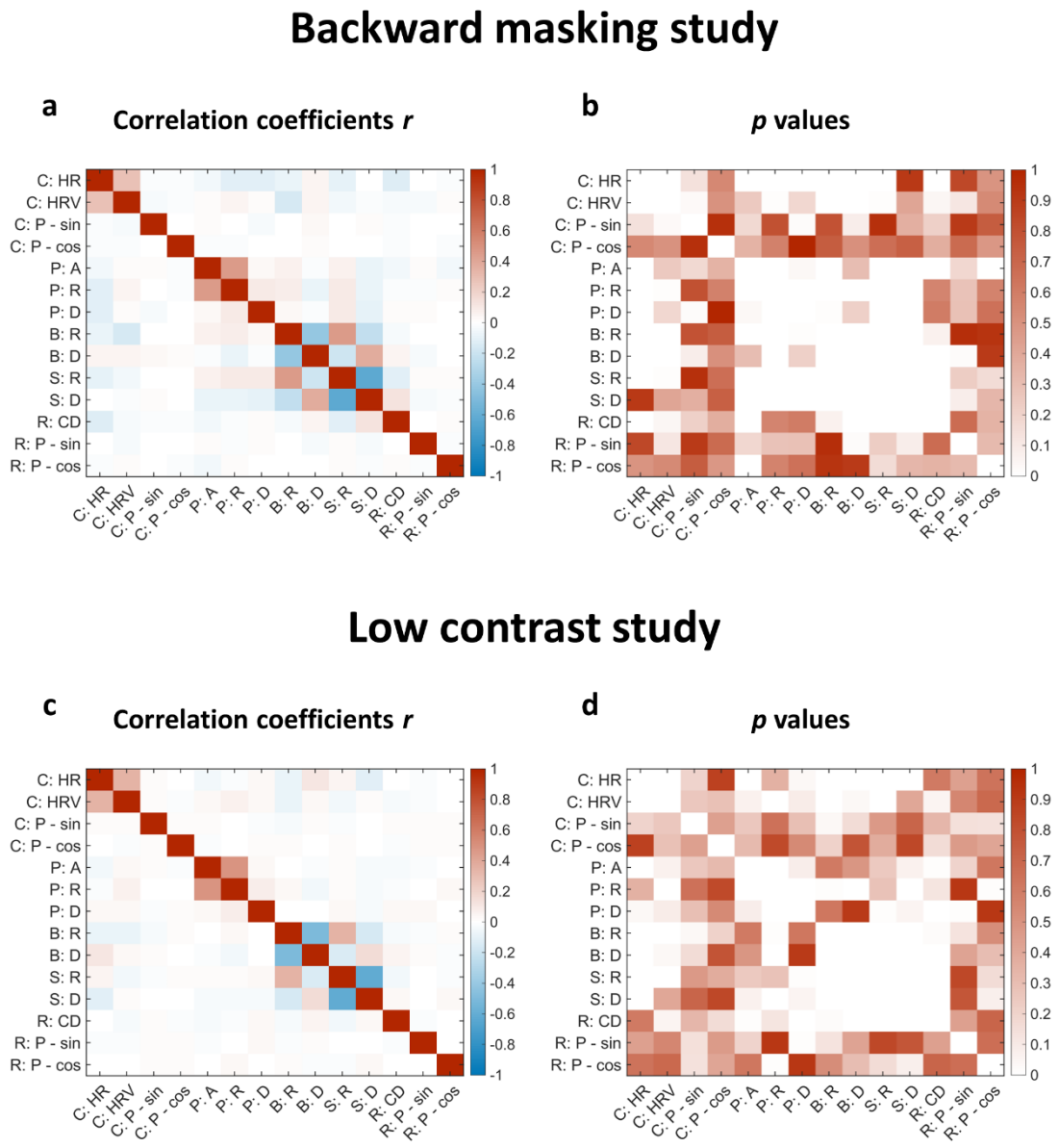


Figure 5.3 *– Correlation between parameters of different physiological signals.* ***a*** *Correlation coefficients (backward masking study).* ***b*** *Significance values (p-values) of the correlation (backward masking study).* ***c*** *Correlation coefficients (low contrast study).* ***d*** *Significance values (p-values) of the correlation (low contrast study). C: HR – Heart rate; C: HRV – Heart rate variation; C: P – sin – sine component of the cardiac phase; C: P – cos – cosine component of the cardiac phase; P: A –pupil size; P: R – relative pupil size; P: D – average pupil derivative; B: R – blinks rate; B: D – distance of the last blink; S: R – saccades rate; S: D – distance of the last saccade; R: CD – respiratory cycle duration; R: P – sin – sine component of the respiratory phase; R: P – cos – cosine component of the respiratory phase.*

5.2 Classifiers within-participant

In this section we will only present the analyses carried out within-participant. It should be noted that, at this stage, the permutation test described in the methods section had not been applied yet. That said, in these analyses, only simple *t*-tests were used. We used the one-sample *t*-test to check whether the performance was statistically greater than 0.5, when using AUC as a metric, and greater than 50%, when using accuracy as a metric. To study whether the incorporation of different physiological signals generated a significant increase in performance, we used a paired-sample *t*-test.

5.2.1 Study of each physiological signal independently

Evaluation of the ability of the classifier to predict trials accuracy (backward masking study) and stimulus recognition (low contrast study) using each physiological signal individually. For this, we employed the SVM algorithm.

	Backward masking study		**Low contrast study**	
	AUC (*p*-value)	**Balanced accuracy** (*p*-value)	**AUC** (*p*-value)	**Balanced accuracy** (*p*-value)
ECG	0.806	0.930	0.210	0.455
Respiration	0.821	0.398	0.716	0.565
Pupil	0.688	0.427	0.002	0.019
Blinks	0.610	0.235	0.999	0.716
Saccades	0.043	0.119	0.187	0.509

Table 24 *– P-values resulting from the application of the one-sample t-test to verify whether the performance is statistically greater than 0.5, when using AUC as a metric, and statistically greater than 50%, when using balanced accuracy as a metric.*

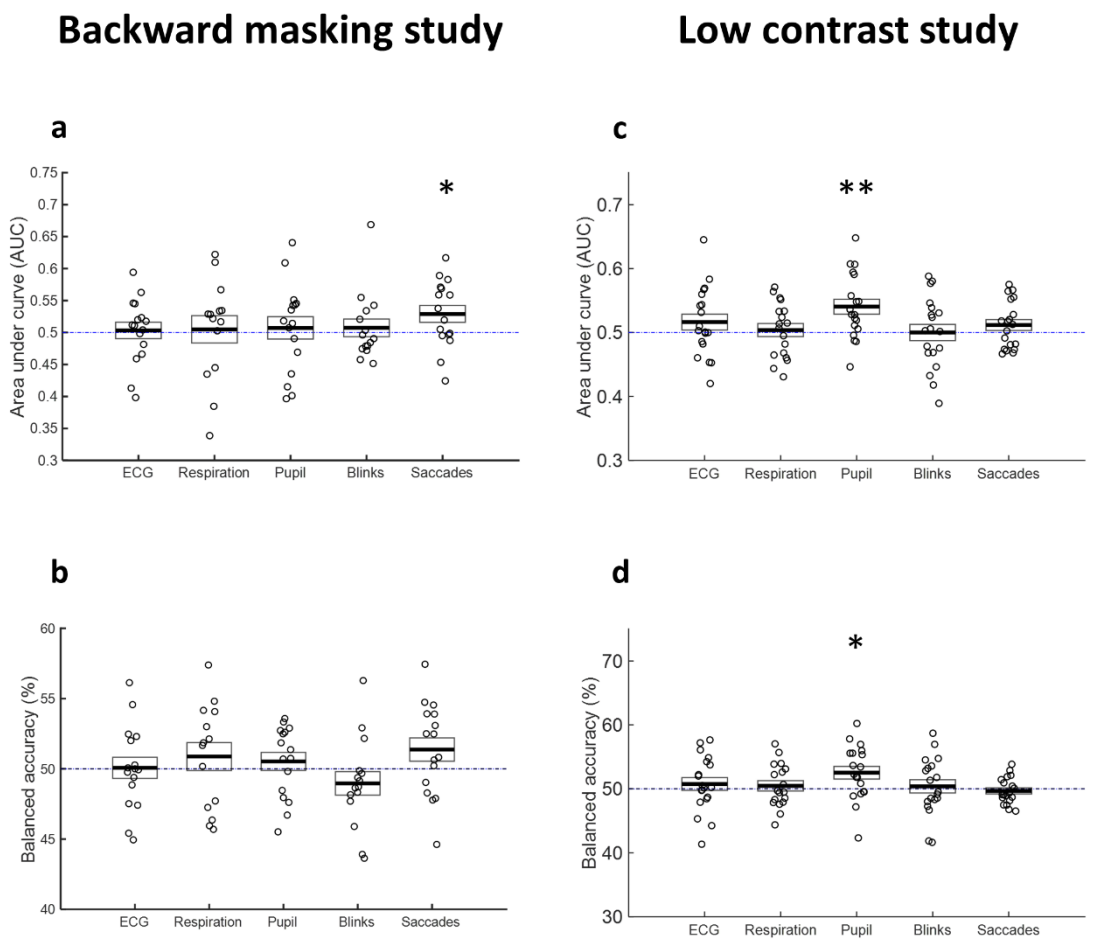


Figure 5.4 *– The performance of the classifier was evaluated for each physiological signal within each participant.* ***a*** *AUC using a SVM model to predict trial performance (backward masking study).* ***b*** *Balanced accuracy using a SVM model to predict trial performance (backward masking study).* ***c*** *AUC using a SVM model to predict stimulus detection (low contrast study).* ***d*** *Balanced accuracy using a SVM model to predict stimulus detection (low contrast study). The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each participant. *: the performance using this physiological signal is statistically better than 0.5. *p<0.05, **p≤0.01, ***p≤0.001*

5.2.2 Combination of each physiological signal with EEG pre-stimulus

This analysis aimed to investigate the influence of pre-stimulus activity on trial detection. To determine whether the performance of the classifier is statistically greater than chance level (0.5 or 50%), we used the one-sample *t*-test. Furthermore, to evaluate whether the inclusion of each pre-stimulus activity significantly modulated the prediction capacity of the classifier and if it resulted in a statistically superior performance compared to using only

neural pre-stimulus activity as input. For these comparisons, we employed the paired sample *t*-test.

	Backward masking study		**Low contrast study**	
	AUC (*p*-value)	**Balanced accuracy** (*p*-value)	**AUC** (*p*-value)	**Balanced accuracy** (*p*-value)
Only EEG pre-stimulus	0.744	0.334	0.023	0.676
EEG + ECG	0.364	0.124	0.003	0.002
EEG + Respiration	0.705	0.746	0.003	0.002
EEG + Pupil	0.888	0.086	0.030	0.030
EEG + Blinks	0.538	0.761	0.030	0.061
EEG + Saccades	0.516	0.084	0.096	0.161
EEG + All measures	0.615	0.797	0.049	0.068

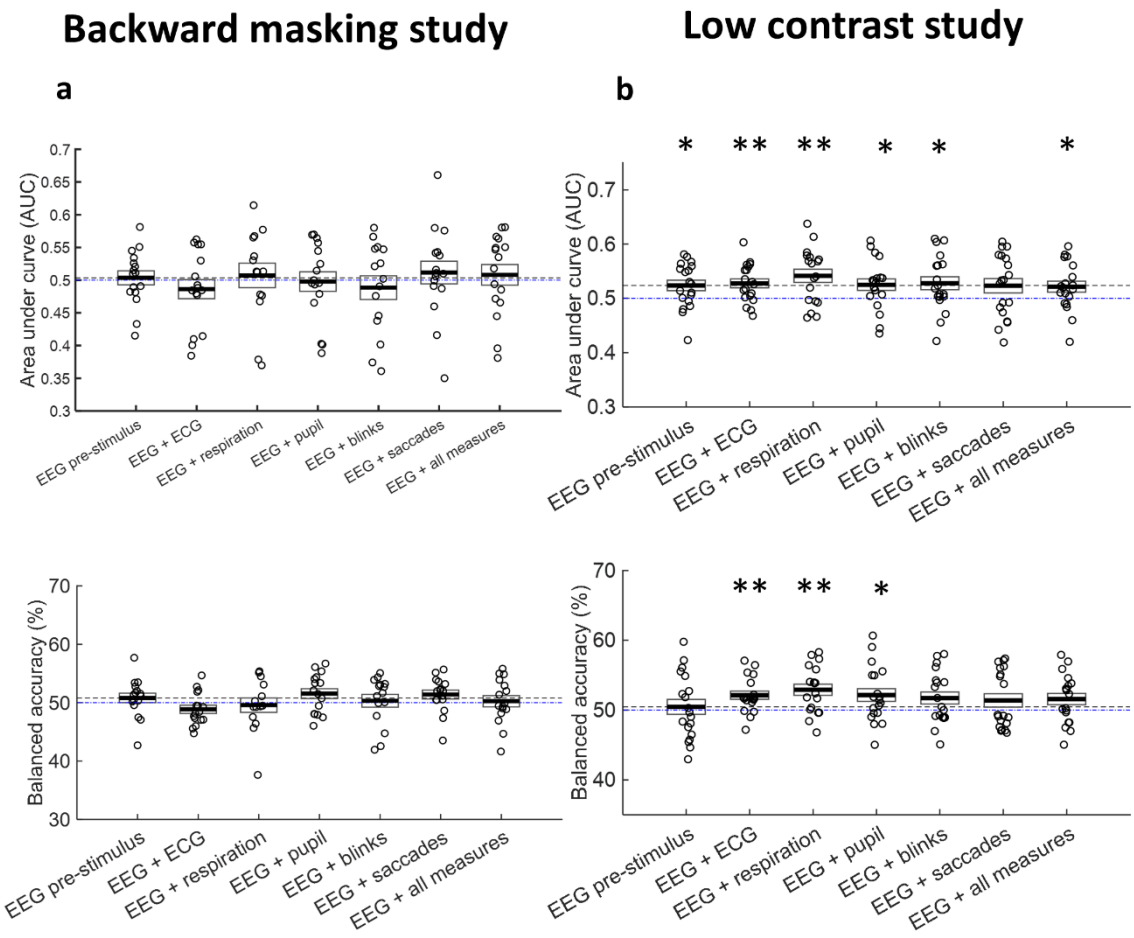
Table 25 *- P-values resulting from the application of the one-sample t-test to verify whether the performance is statistically greater than 0.5, when using AUC as a metric, and statistically greater than 50%, when using accuracy as a metric.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
EEG + ECG	0.312	0.640
EEG + Respiration	0.559	0.196
EEG + Pupil	0.699	0.887
EEG + Blinks	0.394	0.626
EEG + Saccades	0.652	0.977
EEG + All measures	0.805	0.872

Table 26 *– P-values resulting from the application of the paired-sample t-test to verify whether the incorporation of each pre-stimulus activity statistically affects the performance of the classifier. To this analysis, we used only the AUC metric.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
EEG + ECG	0.077	0.108
EEG + Respiration	0.411	0.140
EEG + Pupil	0.499	0.195
EEG + Blinks	0.738	0.294
EEG + Saccades	0.188	0.509
EEG + All measures	0.560	0.436

Table 27 *- P-values resulting from the application of the paired-sample t-test to verify whether the incorporation of each pre-stimulus activity improves statistically the performance of the classifier. To this analysis, we used only the balanced accuracy metric.*



* - The performance of the classifiers is statistically better than 0.5, when using AUC as performance metric, and is statistically better than 50, when using balanced accuracy as performance metric.

Figure 5.5 *– Classifier's performance for each participant using neural pre-stimulus activity combined with the pre-stimulus activity of each physiological signal. The incorporation of the pre-stimulus activity did not cause significant changes in the classifier's performance.* ***a*** *AUC using the multimodal classifier to predict*

trials' accuracy (backward masking study). ***b*** *Balanced accuracy using the multimodal classifier to predict trials' accuracy (backward masking study).* ***c*** *AUC using the multimodal classifier to predict stimulus recognition (low contrast study).* ***d*** *Balanced accuracy using the multimodal classifier to predict stimulus recognition (low contrast study). The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each participant. *: the performance of the classifier is statistically better than 0.5, when using AUC as metric, as is statistically better than 50, when using balanced accuracy. *p<0.05, **p≤0.01, ***p≤0.001*

5.2.3 Combination of neural and physiological pre-stimulus activity with ERPs

This analysis consisted in studying the influence of pre-stimulus activity on the decoding of the stimulus. To determine whether the performance of the classifier is statistically greater than chance level (0.5 or 50%), the one-sample *t*-test was employed. Additionally, to evaluate whether the incorporation of each pre-stimulus activity significantly modulated the discrimination capacity of the classifier and whether it produced a statistically superior performance when using only ERPs as input. To make these comparisons, we utilized the paired-sample *t*-test.

	Backward masking study		**Low contrast study**	
	AUC (*p*-value)	**Accuracy** (*p*-value)	**AUC** (*p*-value)	**Accuracy** (*p*-value)
ERP	<0.001	<0.001	0.001	0.002
ERP + EEG pre-stimulus	<0.001	<0.001	<0.001	0.003
ERP + ECG	<0.001	<0.001	0.002	0.007
ERP + Respiration	<0.001	<0.001	0.002	<0.001
ERP + Pupil	<0.001	<0.001	<0.001	0.001
ERP + Blinks	<0.001	<0.001	<0.001	<0.001
ERP + Saccades	<0.001	<0.001	<0.001	<0.001
ERP + All measures	<0.001	<0.001	<0.001	0.003

Table 28 *– P-values resulting from the application of the one-sample t-test to verify whether the performance is statistically greater than 0.5, when using AUC as a metric, and statistically greater than 50%, when using accuracy as a metric.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
ERP + EEG pre-stimulus	0.922	0.628
ERP + ECG	0.797	0.465
ERP + Respiration	0.245	0.494
ERP + Pupil	0.814	0.611
ERP + Blinks	0.839	0.328
ERP + Saccades	0.715	0.290
ERP + All measures	0.946	0.336

Table 29 *- P-values resulting from the application of the paired-sample t-test to verify whether the incorporation of each pre-stimulus activity improves statistically the performance of the classifier. To this analysis, we used only the AUC metric.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
ERP + EEG pre-stimulus	0.807	0.613
ERP + ECG	0.591	0.260
ERP + Respiration	0.749	0.798
ERP + Pupil	0.390	0.653
ERP + Blinks	0.990	0.320
ERP + Saccades	0.592	0.428
ERP + All measures	0.492	0.441

Table 30 *- P-values resulting from the application of the paired-sample t-test to verify whether the incorporation of each pre-stimulus activity improves statistically the performance of the classifier. To this analysis, we used only the accuracy metric.*

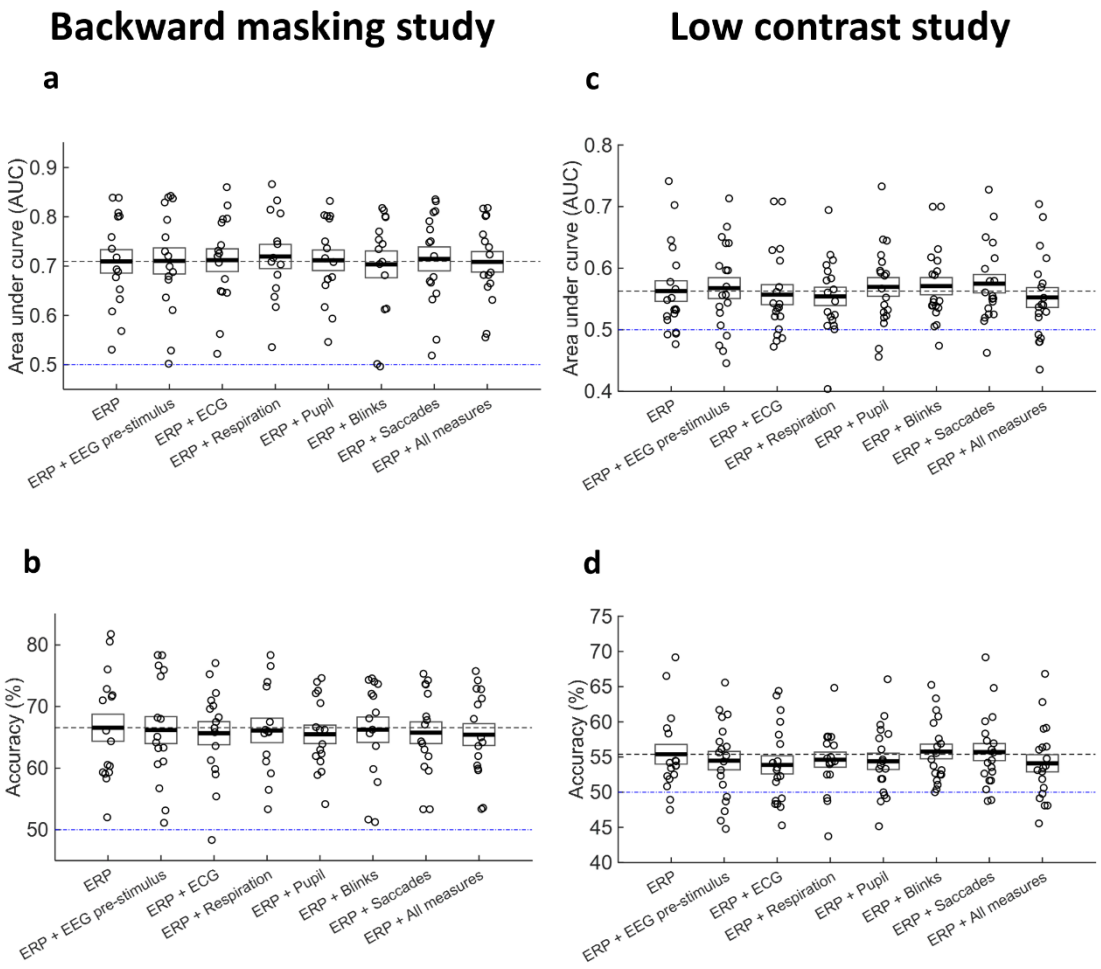


Figure 5.6 *- Classifier's performance for each participant using ERPs combined with the pre-stimulus activity of each physiological signal. The incorporation of the pre-stimulus activity did not cause significant changes in the classifier's performance.* ***a*** *Area under curve using the multimodal classifier to discriminate trials category (backward masking study).* ***b*** *Accuracy using the multimodal classifier to discriminate trials (backward masking study).* ***c*** *Area under curve using the multimodal classifier to discriminate trials category (low contrast study).* ***d*** *Accuracy using the multimodal classifier to discriminate trials category (low contrast study). The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each participant.*

5.3 Classifiers – All trials

In the body of the thesis are the analyses carried out using the AUC as a performance metric. Thus, this section will only contain results using accuracy as a performance metric. As mentioned, it was from these analyses that we introduced the permutation tests.

Then, to analyze the influence of each activity parameter of the different physiological signals, we used the paired-sample t-test.

Finally, we combined pre-stimulus activity with neural activity. To verify whether the incorporation of the pre-stimulus activity produces significant changes in the performance

of the classifier, we used the paired-sample *t*-test. Finally, to analyze the influence of each measure of the different measures extracted from each signal, we also used the paired-sample *t*-test.

5.3.1 Study of each physiological signal independently

Evaluation of the ability of the classifier to predict trials accuracy (backward masking study) and stimulus recognition (low contrast study) using the combination of all parameters of a given physiological signal and each parameter separately. For this, we employed the SVM algorithm.

5.3.1.1 Cardiac activity

	Backward masking study	**Low contrast study**
All cardiac measures	15%	35%
Heart rate	95%	60%
Heart rate variation	5%	40%
Phase	15%	30%

Table 31 *–Percentage of classifiers using cardiac measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
Heart rate	<0.001	0.400
Heart rate variation	0.720	0.309
Phase	0.100	0.871

Table 32 *– Comparison between the use of each cardiac measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all cardiac measures were combined.*

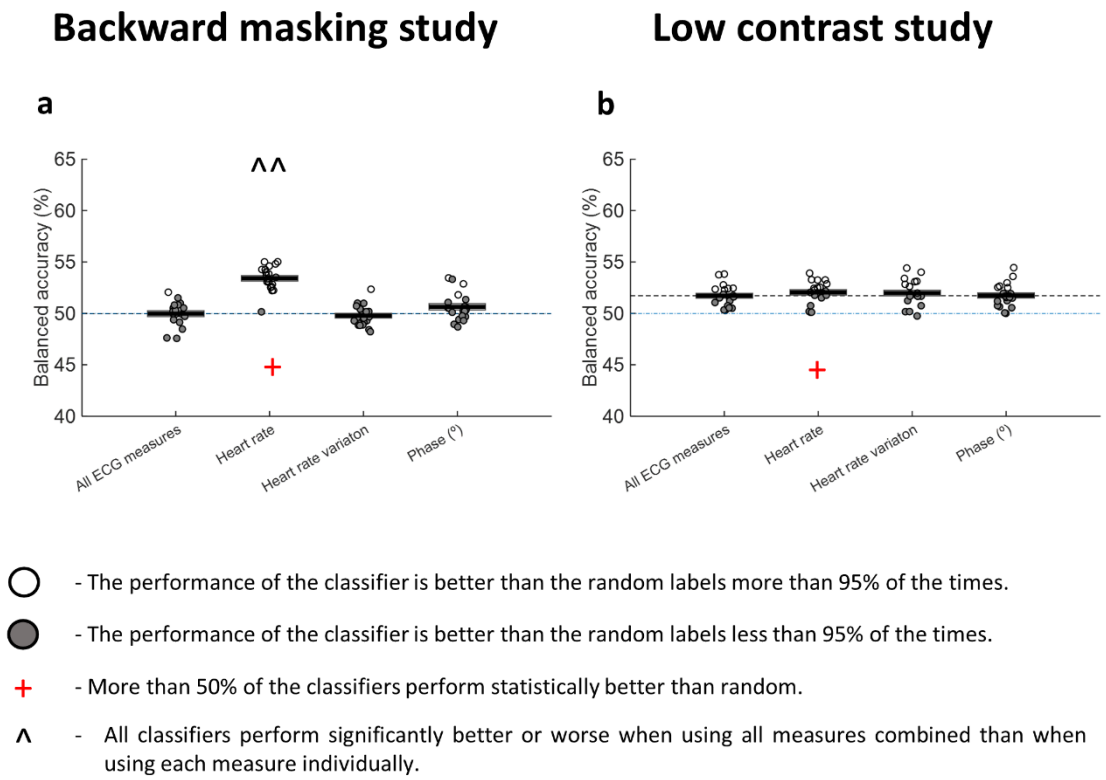


Figure 5.7 *– Classifiers' performance when using all cardiac measures and when using each measure separately.* ***A*** *Balanced accuracy using a SVM model in the backward masking study.* ***b*** *Balanced accuracy using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all cardiac measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

5.3.1.2 Respiratory activity

	Backward masking study	**Low contrast study**
All respiratory measures	15%	100%
Respiratory cycle curation	80%	95%
Phase	70%	100%

Table 33 *–Percentage of classifiers using respiratory measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (p-value)	Low contrast study (p-value)
Respiratory cycle duration	0.002	0.100
Phase	0.003	0.600

Table 34 *– Comparison between the use of each respiratory measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all respiratory measures were combined.*

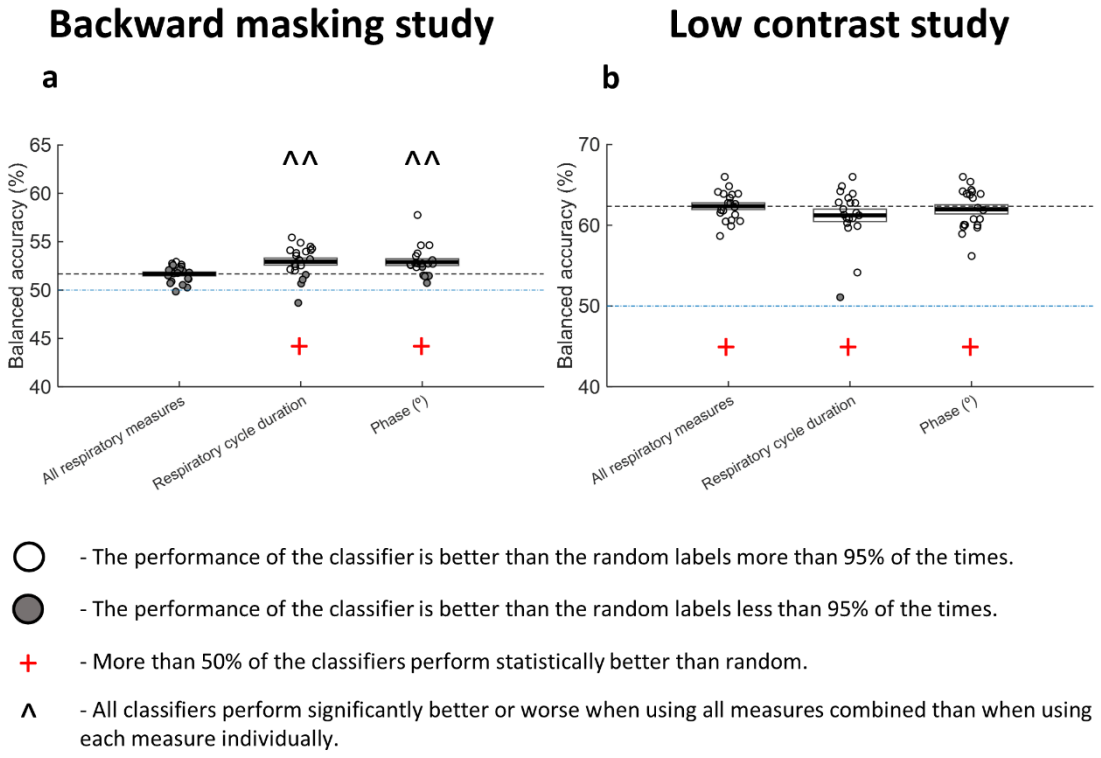


Figure 5.8 *– Classifiers' performance when using all respiratory measures and when using each measure separately.* ***a*** *Balanced accuracy using a SVM model in the backward masking study.* ***b*** *Balanced accuracy using a SVM model the low contrast study. The 1st bar shows the results for when the combination of all respiratory measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The horizontal black line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

5.3.1.3 Pupillary response

	Backward masking study	Low contrast study
All pupil measures	30%	95%
Pupil size	20%	90%
Relative pupil size	40%	100%
Average pupil derivative	15%	95%

Table 35 *- Percentage of classifiers using pupil measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (p-value)	Low contrast study (p-value)
Pupil size	0.838	0.306
Relative pupil size	0.800	0.332
Average pupil derivative	0.401	0.400

Table 36 *– Comparison between the use of each pupil measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all pupil measures were combined.*

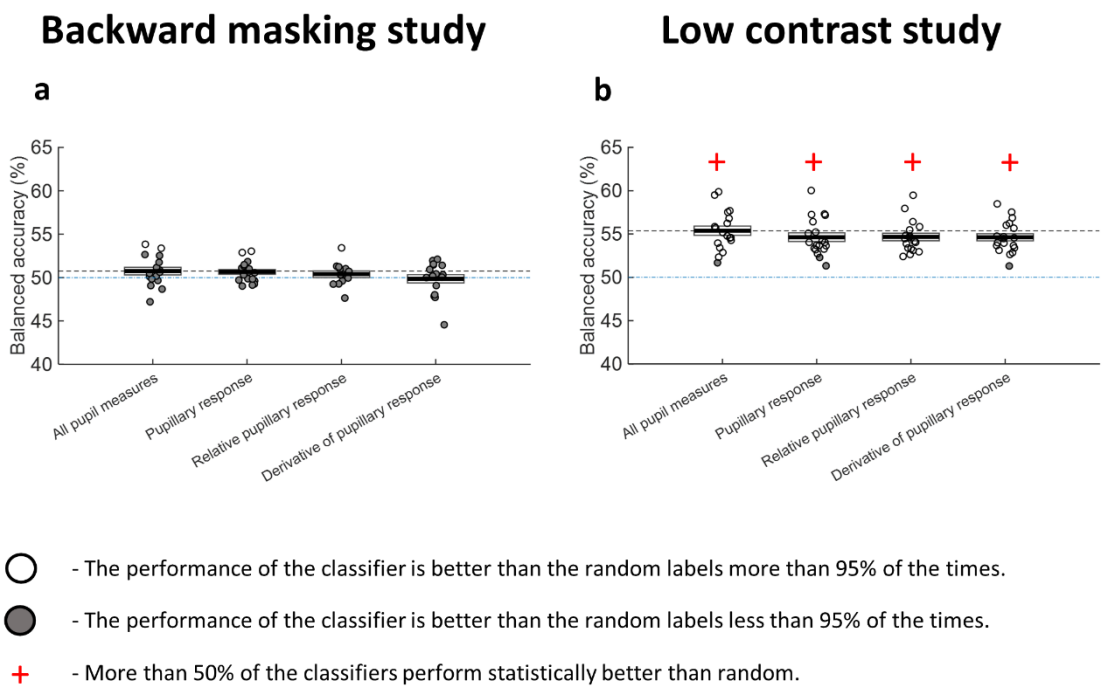


Figure 5.9 *– Classifiers' performance when using all pupil measures and when using each measure separately.* ***a*** *Balanced accuracy using a SVM model in the backward masking study.* ***b*** *Balanced accuracy using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all pupil measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random.*

5.3.1.4 Blinking activity

	Backward masking study	Low contrast study
All blink measures	25%	60%
Blink rate	100%	75%
Distance of the last blink	30%	55%

Table 37 *- Percentage of classifiers using blink measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
Blink rate	<0.001	0.903
Distance of the last blink	0.684	0.829

Table 38 *– Comparison between the use of each blinking measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all blinking measures were combined.*

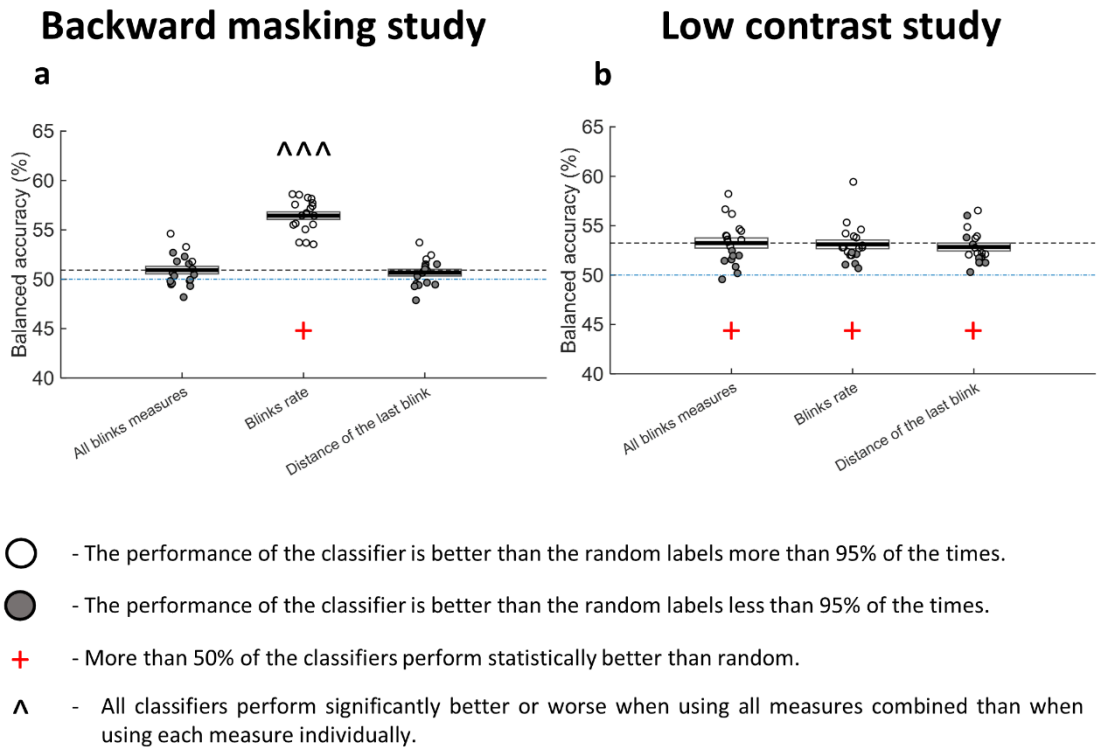


Figure 5.10 *– Classifiers' performance when using all blinking measures and when using each measure separately.* ***a*** *Balanced accuracy using a SVM model in the backward masking study.* ***b*** *Balanced accuracy using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all blinking measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

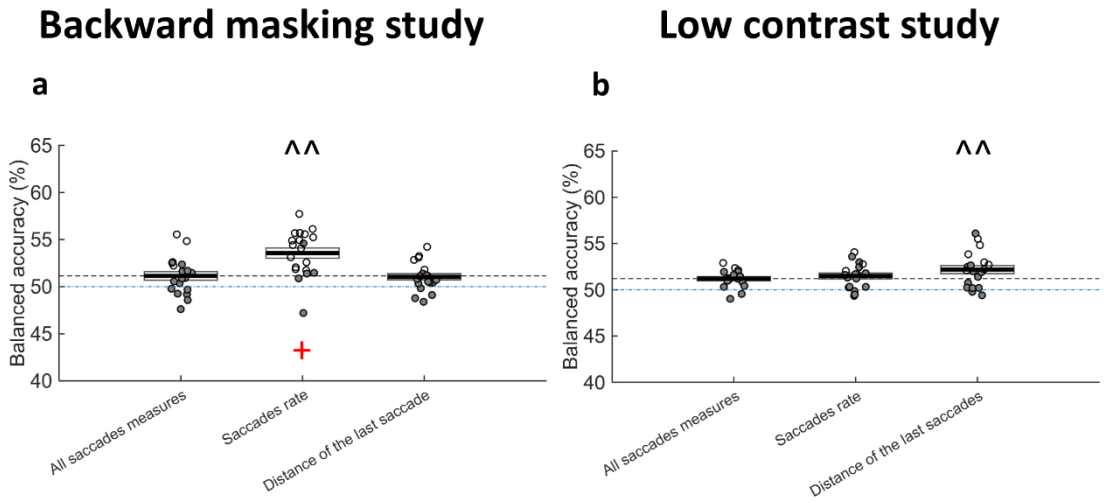
5.3.1.5 Saccadic activity

	Backward masking study	Low contrast study
All saccade measures	25%	25%
Saccades rate	75%	25%
Distance of the last saccade	25%	40%

Table 39 *- Percentage of classifiers using saccade measures as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
Saccades rate	0.002	0.470
Distance of the last saccade	0.824	0.003

Table 40 *– Comparison between the use of each saccade measure separately and the combination of all measures. We used the paired-sample t-test to evaluate the contribution of each measure when all saccades' measures were combined.*



○ - The performance of the classifier is better than the random labels more than 95% of the times.

● - The performance of the classifier is better than the random labels less than 95% of the times.

\+ - More than 50% of the classifiers perform statistically better than random.

^ - All classifiers perform significantly better or worse when using all measures combined than when using each measure individually.

Figure 5.11 *– Classifiers' performance when using all saccades measures and when using each measure separately.* ***a*** *Balanced accuracy using a SVM model in the backward masking study.* ***b*** *Balanced accuracy using a SVM model in the low contrast study. The 1st bar shows the results for when the combination of all saccades measures is used. The remaining bars concern the performance of the classifiers using each of the measures separately. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this measure as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this measure as input, it results in a performance that is statistically different to the performance produced when using all measures combined. ^p<0.05; ^^p<0.01; ^^^p<0.001.*

5.3.2 Combination of each physiological signal with EEG pre-stimulus

We started this analysis by evaluating the capacity of the multimodal classifier to extract activity patterns from the physiological signals and predict visual performance. We employed the permutation tests to evaluate whether the classifier's behavior is better than the random behavior. Then, we compared the classifier's performance when using each physiological signal individually and when combining all physiological signals.

	Backward masking study	Low contrast study
ECG	13.33%	40%
Respiration	13.33%	20%
Pupil	6.67%	80%
Blinks	33.33%	13.33%
Saccades	40%	53.33%
All measures	33.33%	80%

Table 41 *– Percentage of classifiers using only physiological data as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (*p*-value)	Low contrast study (*p*-value)
ECG	0.003	0.172
Respiration	0.002	0.004
Pupil	0.018	0.392
Blinks	0.943	0.481
Saccades	0.498	0.204

Table 42 *– Comparison between the performance achieved using each physiological signal individually and the performance attained through the combination of all physiological signals. We used the paired-sample t-test to evaluate the contribution of each signal when all physiological signals were combined.*

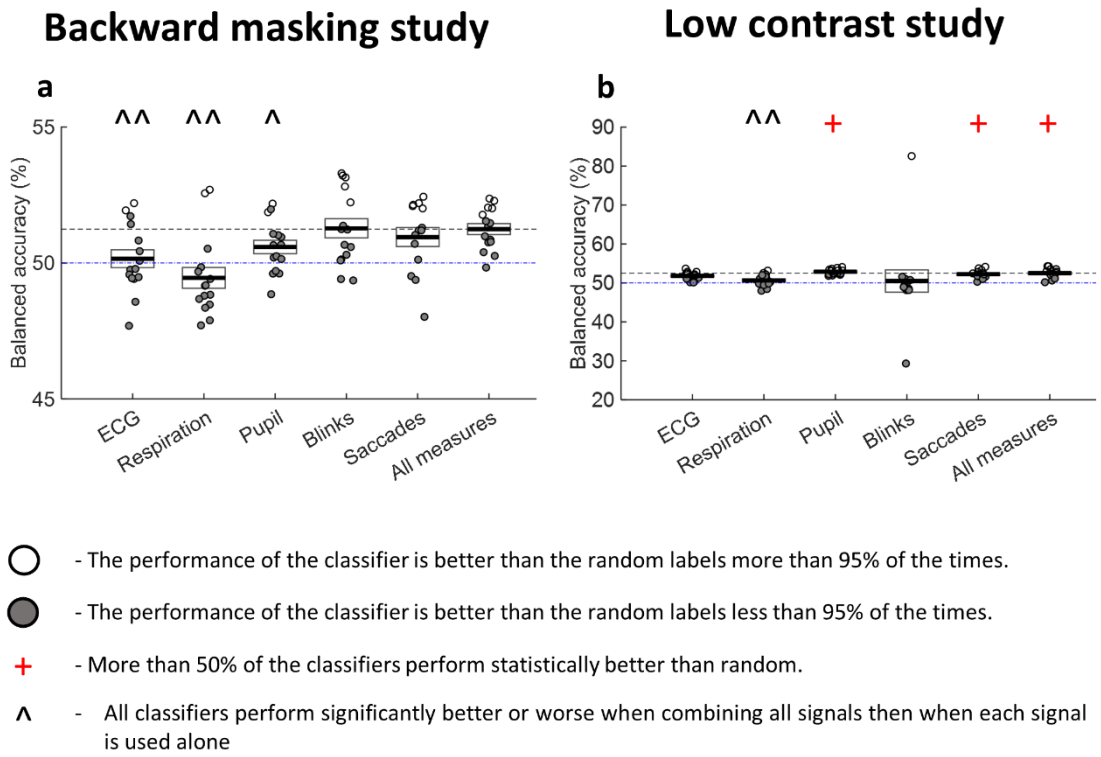


Figure 5.12 *– Multimodal classifiers' performance when using physiological pre-stimulus activity.* ***A*** *Balanced accuracy using a CNN model to predict trial performance (backward masking study).* ***b*** *Balanced accuracy using a CNN model to predict stimulus detection (low contrast study). The last bar shows the results for when the combination of all physiological signals is used. The remaining bars concern the performance of the classifiers using each physiological signal individually. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this pre-stimulus activity as input, it results in a performance that is statistically different to the performance produced when all physiological signals are used. ^p<0.05, ^^p≤0.01, ^^^p≤0.001*

We then combined body and brain pre-stimulus activity. This analysis consisted in studying the influence of pre-stimulus activity in trial's detection. We employed the permutation tests to evaluate whether the classifier's behavior is better than the random behavior. Additionally, to evaluate whether the incorporation of each pre-stimulus activity significantly modulated the prediction capacity of the classifier and whether it produced a statistically superior performance when using only neural pre-stimulus activity as input. To make these comparisons, we utilized the paired sample *t*-test.

	Backward masking study	**Low contrast study**
Only EEG pre-stimulus	86.67%	86.67%
EEG + ECG	60%	100%
EEG + Respiration	73.33%	100%
EEG + Pupil	73.33%	100%
EEG + Blinks	93.33%	100%
EEG + Saccades	66.67%	100%
EEG + All measures	60%	100%

Table 43 *– Percentage of classifiers using the physiological data and neural activity as features that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (*p*-value)	**Low contrast study** (*p*-value)
EEG + ECG	0.021	0.730
EEG + Respiration	0.040	0.586
EEG + Pupil	0.058	0.921
EEG + Blinks	0.021	0.714
EEG + Saccades	0.015	0.418
EEG + All measures	0.008	0.814

Table 44 *– Comparison between the performance achieved using only neural activity and the performance attained through the combination of neural activity with different physiological signals. We used the paired-sample t-test to evaluate whether each physiological signal contains supplementary information that can be extracted by the classifier.*

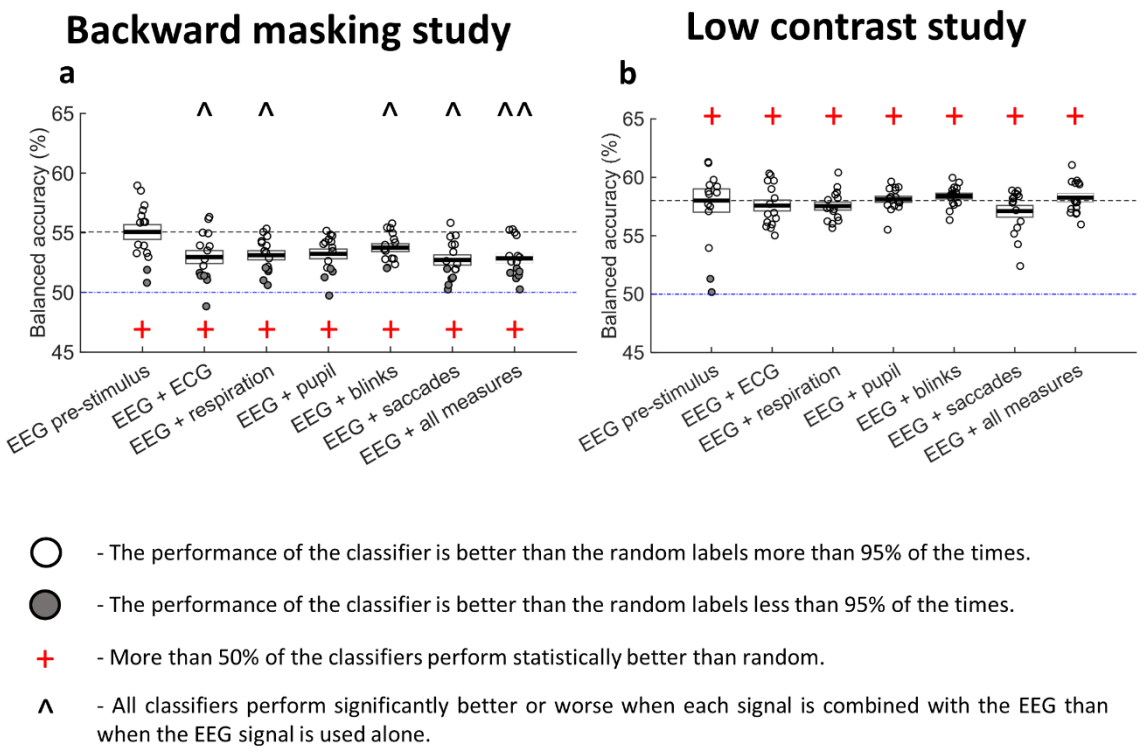


Figure 5.13 *– Classifiers' performance when using neural and physiological pre-stimulus activity.* ***A*** *Balanced accuracy using a CNN model to predict trial performance (backward masking study).* ***b*** *Balanced accuracy using a CNN model to predict stimulus detection (low contrast study). The 1st bar shows the results for when the neural activity pre-stimulus is used. The remaining bars concern the performance of the classifiers combining each physiological signal pre-stimulus with neural activity. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.* ***^****: When utilizing this pre-stimulus activity as input, it results in a performance that is statistically different to the performance produced when only neural pre-stimulus activity is used. ^p<0.05, ^^p≤0.01, ^^^p≤0.001*

5.3.3 Combination of neural and physiological pre-stimulus activity with ERPs

This analysis consisted in studying the influence of pre-stimulus activity on the decoding of the stimulus. We, again, employed the permutation tests to evaluate whether the classifier's behavior is better than the random behavior. Additionally, to evaluate whether the incorporation of each pre-stimulus activity significantly modulated the discrimination capacity of the classifier and whether it produced a statistically superior performance when using only ERPs as input. To make these comparisons, we utilized the paired sample *t*-test.

	Backward masking study	Low contrast study
ERP	100%	93.33%
ERP + EEG pre-stimulus	100%	93.33%
ERP + ECG	100%	86.67%
ERP + Respiration	100%	86.67%
ERP + Pupil	100%	100%
ERP + Blinks	100%	100%
ERP + Saccades	100%	100%
ERP + All measures	100%	93.33%

Table 45 *- Percentage of classifiers using the combination of body and brain pre-stimulus activity with ERPs that perform better than random. We used the permutation test to study if the classifiers had a behavior better than random. The permutation test results give the percentage of classifiers that perform better than random as a metric. That said, to define whether a given measure allows statistical prediction, a threshold of 50% was employed. This signifies that we consider successful classification when a minimum of 50% of the classifiers exhibit a performance surpassing random chance.*

	Backward masking study (p-value)	**Low contrast study** (p-value)
ERP + EEG pre-stimulus	0.104	0.008
ERP + ECG	0.571	0.384
ERP + Respiration	0.918	0.325
ERP + Pupil	0.770	0.774
ERP + Blinks	0.204	0.048
ERP + Saccades	0.450	0.089
ERP + All measures	0.328	0.030

Table 46 *- Comparison between the performance achieved using only ERPs and the performance attained through the combination of ERPs with different pre-stimulus physiological and neural activity. We used the paired-sample t-test to evaluate whether each pre-stimulus activity modulates how stimulus is decoded.*

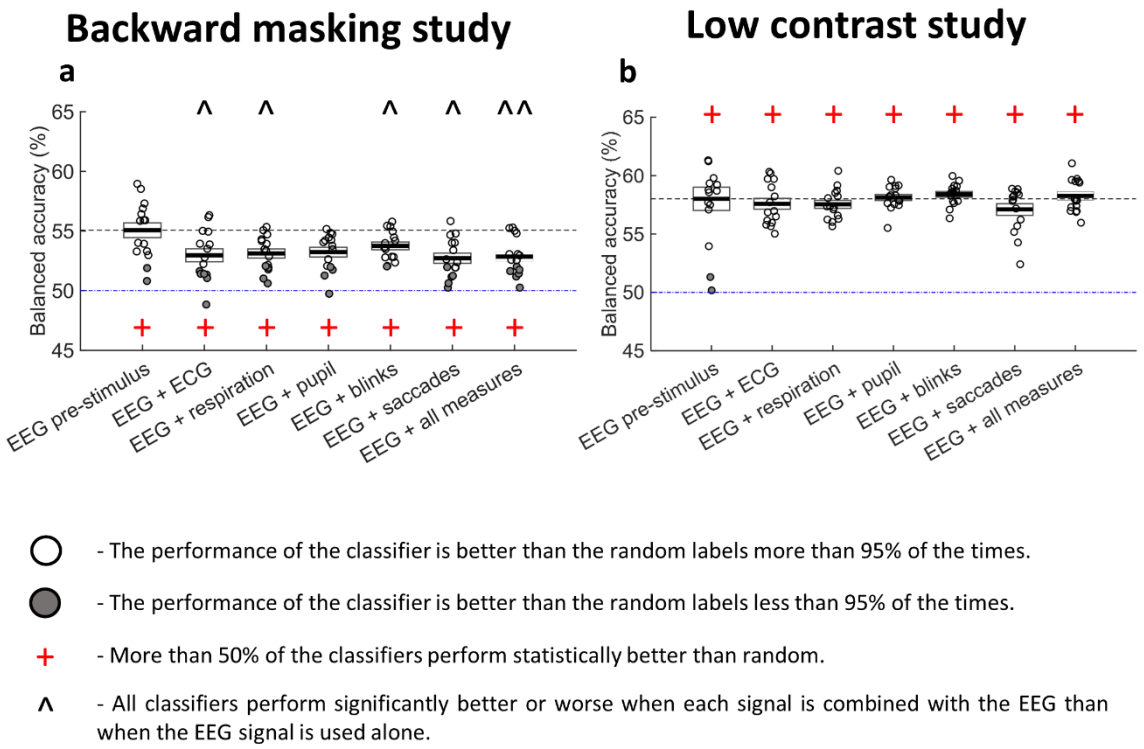


Figure 5.14 *– Classifiers' performance using ERPs combined with the pre-stimulus activity.* ***a*** *Accuracy using a CNN model to discriminate stimulus category (backward masking study).* ***b*** *Accuracy using a CNN model to discriminate stimulus category (low contrast study). The 1st bar shows the results for when only ERPs are used. The remaining bars concern the performance of the classifiers combining pre-stimulus physiological and neural activity with the ERPs. The black horizontal line is the average performance and the rectangle represents ± standard error of the mean. Each circle within the representation denotes each individual instance in which the classifier was executed.* ***+****: Using this pre-stimulus activity as an input, more than 50% of the classifiers present a behavior superior to random.*

References

1. Li Y, Ward MJ, Richardson RM, G'Sell M, Ghuman AS. Endogenous activity modulates stimulus and circuit-specific neural tuning and predicts perceptual behavior. Nat Commun [Internet]. 2020;11(1):4014. Available from: https://doi.org/10.1038/s41467-020-17729-w

2. Walker BB, Sandman CA. Human visual evoked responses are related to heart rate. J Comp Physiol Psychol. 1979;93:717–29.

3. van der Wel P, van Steenbergen H. Pupil dilation as an index of effort in cognitive control tasks: A review. Psychon Bull Rev [Internet]. 2018;25(6):2005–15. Available from: https://doi.org/10.3758/s13423-018-1432-y

4. Perl O, Ravia A, Rubinson M, Eisen A, Soroka T, Mor N, et al. Human non-olfactory cognition phase-locked with inhalation. Nat Hum Behav [Internet]. 2019;3(5):501–12. Available from: https://doi.org/10.1038/s41562-019-0556-z

5. Samaha J, Iemi L, Haegens S, Busch N. Spontaneous brain oscillations and visual perceptual decision making. J Vis [Internet]. 2020 Oct 20;20(11):1729. Available from: https://doi.org/10.1167/jov.20.11.1729

6. Harris KD, Thiele A. Cortical state and attention. Nat Rev Neurosci [Internet]. 2011;12(9):509–23. Available from: https://doi.org/10.1038/nrn3084

7. Cardiovascular Psychophysiology: a Perspective. By P. A. Obrist. (Pp. 236; illustrated; $22.50.) Plenum Press: New York. 1981. Psychol Med [Internet]. 2009/07/09. 1982;12(1):218–218. Available from: https://www.cambridge.org/core/article/cardiovascular-psychophysiology-a-perspective-by-p-a-obrist-pp-236-illustrated-2250-plenum-press-new-york-1981/319DA6F25E9EEB05837D398D88F458F9

8. Roelofs K. Freeze for action: neurobiological mechanisms in animal and human freezing. Philosophical Transactions of the Royal Society B: Biological Sciences. 2017;372.

9. Ribeiro MJ, Castelo-Branco M. Neural correlates of anticipatory cardiac deceleration and its association with the speed of perceptual decision-making, in young and older adults. Neuroimage [Internet]. 2019;199:521–33. Available from: https://www.sciencedirect.com/science/article/pii/S1053811919304896

10. Podvalny E, Flounders MW, King LE, Holroyd T, He BJ. A dual role of prestimulus spontaneous neural activity in visual object recognition. Nat Commun [Internet]. 2019;10(1):3910. Available from: https://doi.org/10.1038/s41467-019-11877-4

11. Stoica Petre, Moses RL. Spectral analysis of signals. Pearson/Prentice Hall; 2005. 452 p.

12. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa O V. Magnetoencephalography---theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev Mod Phys [Internet]. 1993 Apr 1;65(2):413–97. Available from: https://link.aps.org/doi/10.1103/RevModPhys.65.413

13. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol [Internet]. 1962 Jan 1;160(1):106–54. Available from: https://doi.org/10.1113/jphysiol.1962.sp006837

14. Bruce V, Green PR, Georgeson M. Visual perception : Physiology, psychology and ecology. 4th ed. Psychology Press; 2003.

15. Kolb B, Whishaw IQ. Fundamentals of human neuropsychology. Worth Publishers; 2009.

16. Binder M. Encyclopedia of Neuroscience. 2009.

17. Learn About Diagram Of Visual Pathway | Chegg.com [Internet]. [cited 2023 Jul 20]. Available from: https://www.chegg.com/learn/topic/diagram-of-visual-pathway

18. Huff T, Mahabadi N, Tadi P. Neuroanatomy, Visual Cortex [Internet]. StatPearls Publishing, Treasure Island (FL); 2022. Available from: http://europepmc.org/abstract/MED/29494110

19. Mather G. School of Life Sciences: University of Sussex. University of Sussex. "The Visual Cortex."

20. Purves D, Augustine GJ, Fitzpatrick D, Hall WC, LaMantia AS, McNamara JO, et al. Neuroscience, 3rd ed. In 2004. Available from: https://api.semanticscholar.org/CorpusID:147844717

21. Mishkin M, Ungerleider LG. Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. Behavioural Brain Research [Internet]. 1982;6(1):57–77. Available from: https://www.sciencedirect.com/science/article/pii/016643288290081X

22. Revie L. Visual perception and attention and their neurochemical and microstructural brain correlates in healthy and pathological ageing. Cardiff University; 2021.

23. Groen IIA, Silson EH, Baker CI. Contributions of low- and high-level properties to neural processing of visual scenes in the human brain. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2017 Feb 19;372(1714):20160102. Available from: https://doi.org/10.1098/rstb.2016.0102

24. Gilbert CD, Li W. Top-down influences on visual processing. Nat Rev Neurosci [Internet]. 2013;14(5):350–63. Available from: https://doi.org/10.1038/nrn3476

25. Teufel C, Dakin S, Fletcher P. Object Knowledge Shapes Properties of Early Feature-Detectors by Top-Down Modulation. J Vis [Internet]. 2015 Sep 1;15(12):1029. Available from: https://doi.org/10.1167/15.12.1029

26. Ronchi R, Bernasconi F, Pfeiffer C, Bello-Ruiz J, Kaliuzhna M, Blanke O. Interoceptive signals impact visual processing: Cardiac modulation of visual body perception. Neuroimage [Internet]. 2017;158:176–85. Available from: https://www.sciencedirect.com/science/article/pii/S1053811917305359

27. Becker DE. Fundamentals of Electrocardiography Interpretation. Vol. 53, Anesth Prog. 2006.

28. Guyton AC, Hall JE. Textbook of Medical Physiology. In 1961. Available from: https://api.semanticscholar.org/CorpusID:72200532

29. Kebe M, Gadhafi R, Mohammad B, Sanduleanu M, Saleh H, Al-Qutayri M. Human Vital Signs Detection Methods and Potential Using Radars: A Review. Sensors [Internet]. 2020;20(5). Available from: https://www.mdpi.com/1424-8220/20/5/1454

30. Almeida T. Análise Comparativa de Técnicas de Rastreamento de Marcas Acústicas em Imagens de Ecocardiografia. 2012.

31. Gordan R, Gwathmey JK, Xie LH. Autonomic and endocrine control of cardiovascular function. World J Cardiol. 2015;7 4:204–14.

32. Salomon R, Ronchi R, Dönz J, Bello-Ruiz J, Herbelin B, Faivre N, et al. Insula mediates heartbeat related effects on visual consciousness. Cortex [Internet]. 2018;101:87–95. Available from: https://www.sciencedirect.com/science/article/pii/S0010945218300182

33. Larra MF, Finke JB, Wascher E, Schächinger H. Disentangling sensorimotor and cognitive cardioafferent effects: A cardiac-cycle-time study on spatial stimulus-response compatibility. Sci Rep [Internet]. 2020;10(1):4059. Available from: https://doi.org/10.1038/s41598-020-61068-1

34. Azzalini D, Rebollo I, Tallon-Baudry C. Visceral Signals Shape Brain Dynamics and Cognition. Trends Cogn Sci [Internet]. 2019;23(6):488–509. Available from: https://www.sciencedirect.com/science/article/pii/S1364661319300890

35. Marlin DJ, Deaton CM. 15 - Pulmonary Function Testing. In: McGorum BC, Dixon PM, Robinson NE, Schumacher J, editors. Equine Respiratory Medicine and Surgery [Internet]. Edinburgh: W.B. Saunders; 2007. p. 211–33. Available from: https://www.sciencedirect.com/science/article/pii/B9780702027598500209

36. Zelano C, Jiang H, Zhou G, Arora N, Schuele S, Rosenow J, et al. Nasal Respiration Entrains Human Limbic Oscillations and Modulates Cognitive Function. The Journal of Neuroscience [Internet]. 2016 Dec 7;36(49):12448. Available from: http://www.jneurosci.org/content/36/49/12448.abstract

37. Nilsson DE. The Evolution of Eyes. In: Fabian A, Gibson J, Sheppard M, Weyand S, editors. Vision [Internet]. Cambridge: Cambridge University Press; 2021. p. 5–32. (Darwin College Lectures). Available from: https://www.cambridge.org/core/books/vision/evolution-of-eyes/8946D789C2A0C9A07BC51F100353D5C6

38. Steinhauer SR, Siegle GJ, Condray R, Pless M. Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. International Journal of Psychophysiology [Internet]. 2004;52(1):77–86. Available from: https://www.sciencedirect.com/science/article/pii/S0167876003002721

39. Farnsworth B. Pupillometry 101: What You Need to Know - iMotions [Internet]. [cited 2023 May 8]. Available from: https://imotions.com/blog/learning/best-practice/pupillometry-101/

40. Dark and Bright pupil tracking [Internet]. [cited 2023 May 8]. Available from: https://connect.tobii.com/s/article/What-is-dark-and-bright-pupil-tracking?language=en_US

41. Duchowski AT, Duchowski AT. Eye tracking methodology: Theory and practice. Springer; 2017.

42. Naber M, Nakayama K. Pupil responses to high-level image content. J Vis [Internet]. 2013 May 17;13(6):7. Available from: https://doi.org/10.1167/13.6.7

43. Tsukahara JS, Harrison TL, Engle RW. The relationship between baseline pupil size and intelligence. Cogn Psychol [Internet]. 2016;91:109–23. Available from: https://www.sciencedirect.com/science/article/pii/S0010028516300585

44. Stern J. MSNBC. 1999 [cited 2023 May 8]. Why do we blink? Available from: https://www.nbcnews.com/id/wbna3076704

45. Irwin DE. Where does attention go when you blink? Atten Percept Psychophys [Internet]. 2011 Jul 24 [cited 2023 May 8];73(5):1374–84. Available from: https://link.springer.com/article/10.3758/s13414-011-0111-0

46. Shultz S, Klin A, Jones W. Inhibition of eye blinking reveals subjective perceptions of stimulus salience. Proc Natl Acad Sci U S A. 2011 Dec 27;108(52):21270–5.

47. Abeles D, Amit R, Tal-Perry N, Carrasco M, Yuval-Greenberg S. Oculomotor inhibition precedes temporally expected auditory targets. Nat Commun [Internet]. 2020;11(1):3524. Available from: https://doi.org/10.1038/s41467-020-17158-9

48. Magliacano A, Fiorenza S, Estraneo A, Trojano L. Eye blink rate increases as a function of cognitive load during an auditory oddball paradigm. Neurosci Lett [Internet]. 2020;736:135293. Available from: https://www.sciencedirect.com/science/article/pii/S0304394020305632

49. Paprocki R, Lenskiy A. What does eye-blink rate variability dynamics tell us about cognitive performance? Front Hum Neurosci. 2017 Dec 19;11.

50. Purves D, Augustine GJ, Fitzpatrick D, Katz LC, LaMantia AS, McNamara JO, et al. Types of Eye Movements and Their Functions. 2001 [cited 2023 May 8]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK10991/

51. Termsarasab P, Thammongkolchai T, Rucker JC, Frucht SJ. The diagnostic value of saccades in movement disorder patients: a practical guide and review. Journal of Clinical Movement Disorders 2015 2:1 [Internet]. 2015 Oct 15 [cited 2023 May 8];2(1):1–10. Available from: https://clinicalmovementdisorders.biomedcentral.com/articles/10.1186/s40734-015-0025-4

52. Land M. Encyclopedia Britannica. 2012 [cited 2023 May 8]. Saccade | physiology | Britannica. Available from: https://www.britannica.com/science/saccade

53. Ibbotson M, Krekelberg B. Visual perception and saccadic eye movements. Vol. 21, Current Opinion in Neurobiology. 2011. p. 553–8.

54. SEER Training Modules, Module Name. U. S. National Institutes of Health, National Cancer Institute. <https://training.seer.cancer.gov/>.

55. Kirschstein T, Köhling R. What is the Source of the EEG? Clin EEG Neurosci [Internet]. 2009 Jul 1;40(3):146–9. Available from: https://doi.org/10.1177/155005940904000305

56. Biasiucci A, Franceschiello B, Murray MM. Electroencephalography. Current Biology [Internet]. 2019;29(3):R80–5. Available from: https://www.sciencedirect.com/science/article/pii/S0960982218315513

57. Kumar JS, Bhuvaneswari P. Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study. Procedia Eng [Internet]. 2012;38:2525–36. Available from: https://www.sciencedirect.com/science/article/pii/S1877705812022114

58. Beniczky S, Schomer DL. Electroencephalography: basic biophysical and technological aspects important for clinical applications. Epileptic Disorders. 2020 Dec 1;22(6):697–715.

59. Nuwer MR, Comi G Pietro, Emerson R, Fuglsang-Frederiksen A, Guérit JM, Hinrichs H, et al. IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. Electroencephalogr Clin Neurophysiol [Internet]. 1998;106 3:259–61. Available from: https://api.semanticscholar.org/CorpusID:32033816

60. Zhou YJ, Iemi L, Schoffelen JM, Lange FP de, Haegens S. Alpha oscillations shape sensory representation and perceptual sensitivity. bioRxiv [Internet]. 2021 Jan 1;2021.02.02.429418. Available from: http://biorxiv.org/content/early/2021/07/30/2021.02.02.429418.abstract

61. Brüers S, VanRullen R. Alpha Power Modulates Perception Independently of Endogenous Factors. Front Neurosci. 2018 Apr 25;12.

62. Posner MI, Petersen SE. The Attention System of the Human Brain. Annu Rev Neurosci [Internet]. 1990 Mar 1;13(1):25–42. Available from: https://doi.org/10.1146/annurev.ne.13.030190.000325

63. Oken BS, Salinsky MC, Elsas SM. Vigilance, alertness, or sustained attention: physiological basis and measurement. Clinical Neurophysiology [Internet]. 2006;117(9):1885–901. Available from: https://www.sciencedirect.com/science/article/pii/S1388245706000496

64. Tseng VWS, Abdullah S, Costa JDR, Choudhury T. AlertnessScanner: what do your pupils tell about your alertness. Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services. 2018;

65. Périn B, Godefroy O, Fall S, de Marco G. Alertness in young healthy subjects: An fMRI study of brain region interactivity enhanced by a warning signal. Brain Cogn [Internet]. 2010;72(2):271–81. Available from: https://www.sciencedirect.com/science/article/pii/S027826260900181X

66. Hackley SA. The speeding of voluntary reaction by a warning signal. Psychophysiology [Internet]. 2009;46(2):225–33. Available from: http://europepmc.org/abstract/MED/18811626

67. Walter WG, Cooper R, Aldridge VJ, Mccallum WC, Winter AL. Contingent Negative Variation : An Electric Sign of Sensori-Motor Association and Expectancy in the Human Brain. Nature. 1964;203:380–4.

68. Mannarelli D, Pauletti C, Grippo A, Amantini A, Augugliaro V, Currà A, et al. The Role of the Right Dorsolateral Prefrontal Cortex in Phasic Alertness: Evidence from a Contingent Negative Variation and Repetitive Transcranial Magnetic Stimulation Study. Mangel SC, editor. Neural Plast [Internet]. 2015;2015:410785. Available from: https://doi.org/10.1155/2015/410785

69. Pauletti C, Mannarelli D, Grippo A, Currà A, Locuratolo N, De Lucia MC, et al. Phasic alertness in a cued double-choice reaction time task: A Contingent Negative Variation (CNV) study. Neurosci Lett [Internet]. 2014;581:7–13. Available from: https://www.sciencedirect.com/science/article/pii/S0304394014006429

70. Guo Z, Tan X, Pan Y, Liu X, Zhao G, Wang L, et al. Contingent negative variation during a modified cueing task in simulated driving. PLoS One [Internet]. 2019 Nov 11;14(11):e0224966-. Available from: https://doi.org/10.1371/journal.pone.0224966

71. Sturm W, de Simone A, Krause BJ, Specht K, Hesselmann V, Radermacher I, et al. Functional anatomy of intrinsic alertness: evidence for a fronto-parietal-thalamic-brainstem network in the right hemisphere. Neuropsychologia [Internet]. 1999;37(7):797–805. Available from: http://europepmc.org/abstract/MED/10408647

72. Sturm W, Willmes K. On the Functional Neuroanatomy of Intrinsic and Phasic Alertness. Neuroimage [Internet]. 2001;14(1):S76–84. Available from: https://www.sciencedirect.com/science/article/pii/S1053811901908390

73. Castro-Alamancos MA, Gulati T. Neuromodulators Produce Distinct Activated States in Neocortex. The Journal of Neuroscience [Internet]. 2014 Sep 10;34(37):12353. Available from: http://www.jneurosci.org/content/34/37/12353.abstract

74. Lenartowicz A, Simpson G, Cohen M. Perspective: causes and functional significance of temporal variations in attention control. Front Hum Neurosci [Internet]. 2013;7. Available from: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00381

75. Samuels E, Szabadi E. Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part II: physiological and pharmacological manipulations and pathological alterations of locus coeruleus activity in humans. Curr Neuropharmacol [Internet]. 2008 Sep 11 [cited 2023 Jul 27];6(3):254–85. Available from: https://pubmed.ncbi.nlm.nih.gov/19506724/

76. Parikh V, Bangasser D. Cholinergic Signaling Dynamics and Cognitive Control of Attention. In: Current topics in behavioral neurosciences. 2020.

77. Parikh V, Sarter M. Cholinergic Mediation of Attention. Ann N Y Acad Sci [Internet]. 2008 May 1;1129(1):225–35. Available from: https://doi.org/10.1196/annals.1417.021

78. Mitsushima D. Sex Steroids and Acetylcholine Release in the Hippocampus. Vitam Horm. 2010 Jan 1;82:263–77.

79. Posner MI. Measuring Alertness. Ann N Y Acad Sci [Internet]. 2008 May 1;1129(1):193–9. Available from: https://doi.org/10.1196/annals.1417.011

80. Urai AE, Braun A, Donner TH. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat Commun. 2017 Mar 3;8.

81. Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun. 2016 Nov 8;7.

82. Jennings JR, Van Der Molen MW. Preparation for speeded action as a psychophysiological concept. Psychol Bull. 2005 May;131(3):434–59.

83. Marsland S. Machine Learning: An Algorithmic Perspective. 1st ed. Chapman & Hall/CRC; 2009.

84. Ray S. A Quick Review of Machine Learning Algorithms. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon). 2019;35–9.

85. Shalev-Shwartz S, Ben-David S. Understanding Machine Learning: From Theory to Algorithms [Internet]. Cambridge: Cambridge University Press; 2014. Available from: https://www.cambridge.org/core/books/understanding-machine-learning/3059695661405D25673058E43C8BE2A6

86. Campbell C, Ying Y. Learning with Support Vector Machines. Vol. 5, Synthesis Lectures on Artificial Intelligence and Machine Learning. 2011.

87. Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. Least Squares Support Vector Machines [Internet]. WORLD SCIENTIFIC; 2002. 308 p. Available from: https://doi.org/10.1142/5089

88. Bergstra J, Bardenet R, Bengio Y, Kégl B. Algorithms for Hyper-Parameter Optimization. In: Proceedings of the 24th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc.; 2011. p. 2546–54. (NIPS'11).

89. Fit k-nearest neighbor classifier - MATLAB fitcknn [Internet]. [cited 2023 Jun 12]. Available from: https://www.mathworks.com/help/stats/fitcknn.html?s_tid=doc_ta#bt6cr9l_sep_shared-HyperparameterOptimizationOptions

90. Snoek J, Larochelle H, Adams RP. Practical Bayesian Optimization of Machine Learning Algorithms. In: Pereira F, Burges CJ, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems [Internet]. Curran Associates, Inc.; 2012. Available from: https://proceedings.neurips.cc/paper_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf

91. Garnett R. Bayesian Optimization [Internet]. Cambridge: Cambridge University Press; 2023. Available from: https://www.cambridge.org/core/books/bayesian-optimization/11AED383B208E7F22A4CE1B5BCBADB44

92. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data [Internet]. 2021;8(1):53. Available from: https://doi.org/10.1186/s40537-021-00444-8

93. Patterson J, Gibson A. Deep Learning: A Practitioner's Approach. 1st ed. O'Reilly Media, Inc.; 2017.

94. Shanmugamani R, Moore SM. Deep Learning for Computer Vision: Expert Techniques to Train Advanced Neural Networks Using TensorFlow and Keras [Internet]. Packt Publishing; 2018. Available from: https://books.google.pt/books?id=dgdOswEACAAJ

95. Roy Y, Banville H, Carneiro de Albuquerque IM, Gramfort A, Faubert J. Deep learning-based electroencephalography analysis: a systematic review. 2019.

96. Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. J Neural Eng. 2018 Jul 27;15(5).

97. Options for training deep learning neural network - MATLAB trainingOptions [Internet]. [cited 2023 Jul 20]. Available from: https://www.mathworks.com/help/deeplearning/ref/trainingoptions.html

98. 3.3. Metrics and scoring: quantifying the quality of predictions — scikit-learn 1.2.2 documentation [Internet]. [cited 2023 Jun 12]. Available from: https://scikit-learn.org/stable/modules/model_evaluation.html

99. The Explanation You Need on Binary Classification Metrics | by Andrea D'Agostino | Towards Data Science [Internet]. [cited 2023 Jun 12]. Available from: https://towardsdatascience.com/the-explanation-you-need-on-binary-classification-metrics-321d280b590f

100. What is balanced accuracy? | Statistical Odds & Ends [Internet]. [cited 2023 Jun 12]. Available from: https://statisticaloddsandends.wordpress.com/2020/01/23/what-is-balanced-accuracy/

101. N. G. Sousa, "Body-brain interactions in visual perception," Master's thesis, Retrieved from http://hdl.handle.net/10316/99376, 2022.

102. Stigliani A, Weiner KS, Grill-Spector K. Temporal processing capacity in high-level visual cortex is domain specific. Journal of Neuroscience. 2015;35(36):12412–24.

103. Epstein R, Kanwisher N. A cortical representation of the local visual environment. Nature [Internet]. 1998;392(6676):598–601. Available from: https://doi.org/10.1038/33402

104. Kanwisher N, Yovel G. The fusiform face area: a cortical region specialized for the perception of faces. Philosophical Transactions of the Royal Society B: Biological Sciences [Internet]. 2006 Nov 8;361(1476):2109–28. Available from: https://doi.org/10.1098/rstb.2006.1934

105. Habib M, Sirigu A. Pure Topographical Disorientation: A Definition and Anatomical Basis. Cortex [Internet]. 1987;23(1):73–85. Available from: https://www.sciencedirect.com/science/article/pii/S0010945287800205

106. Kleiner M, Brainard DH, Pelli D, Ingling A, Murray R, Broussard C. What's new in Psychtoolbox-3. Perception. 2007 Jan 1;36:1–16.

107. Oliveira, Guilherme António Duarte, "Decision in visual perception: cortical mechanisms in the interpretation of ambiguous visual stimuli", Master's thesis, Retrieved from http://hdl.handle.net/10316/103126, 2022.

108. Prins N, others. Psychophysics: a practical introduction. Academic Press; 2016.

109. Prins N, Kingdom FAA. Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. Front Psychol. 2018;9:1250.

110. Santos, Manuel da Silva, "Brain signals of perceptual inference: the role of oscillations in interpreting ambiguous stimuli", Master's thesis, Retrieved from http://hdl.handle.net/10316/90113, 2020.

111. Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. Percept Psychophys. 2001;63(8):1293–313.

112. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods [Internet]. 2004;134(1):9–21. Available from: https://www.sciencedirect.com/science/article/pii/S0165027003003479

113. Chaumon M, Bishop DVM, Busch NA. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. J Neurosci Methods [Internet]. 2015;250:47–63. Available from: https://www.sciencedirect.com/science/article/pii/S0165027015000928

114. Crameri F, Shephard GE, Heron PJ. The misuse of colour in science communication. Nature Communications 2020 11:1 [Internet]. 2020 Oct 28 [cited 2023 Jun 19];11(1):1–10. Available from: https://www.nature.com/articles/s41467-020-19160-7

115. Medeiros, Rogério Manuel Pedro, "Brain-Heart Interaction: The Effect of Cardiac activity on Visual Perception", Master's thesis, Retrieved from http://hdl.handle.net/10316/102953, 2022.

116. Henriques J, Rocha T, Paredes S, Cabiddu R, Mendes D, Couceiro R, et al. ECG analysis tool for heart failure management and cardiovascular risk assessment. In 2015. Available from: https://api.semanticscholar.org/CorpusID:43866118

117. Flenady T, Dwyer T, Applegarth J. Accurate respiratory rates count: So should you! Australasian Emergency Nursing Journal [Internet]. 2017;20(1):45–7. Available from: https://www.sciencedirect.com/science/article/pii/S157462671630060X

118. EyeLink ® 1000 User Manual Tower, Desktop, LCD Arm, Primate and Long Range Mounts Remote, 2000 Hz and Fiber Optic Camera Upgrades [Internet]. 2005. Available from: http://www.sr-research.com/

119. Abusharha AA. Changes in blink rate and ocular symptoms during different reading tasks. Clin Optom (Auckl). 2017 Nov 20;9:133–8.

120. FAUST O, BAIRY MG. NONLINEAR ANALYSIS OF PHYSIOLOGICAL SIGNALS: A REVIEW. J Mech Med Biol [Internet]. 2012 Aug 13;12(04):1240015. Available from: https://doi.org/10.1142/S0219519412400155

121. Optimize Classifier Fit Using Bayesian Optimization - MATLAB & Simulink [Internet]. [cited 2023 Jun 12]. Available from: https://www.mathworks.com/help/stats/optimize-an-svm-classifier-fit-using-bayesian-optimization.html

122. Al-Daffaie K, Khan S. Logistic regression for circular data. Vol. 1842, AIP Conference Proceedings. 2017. 030022 p.

123. David Groppe. MATLAB Central File Exchange. mult_comp_perm_t1(data,n_perm,tail,alpha_level,mu,reports,seed_state) (https://www.mathworks.com/matlabcentral/fileexchange/29782-mult_comp_perm_t1-data-n_perm-tail-alpha_level-mu-reports-seed_state).

124. Good PI. Permutation, Parametric, and Bootstrap Tests of Hypotheses (Springer Series in Statistics). Berlin, Heidelberg: Springer-Verlag; 2004.

125. Palmer J, Huk AC, Shadlen MN. The effect of stimulus strength on the speed and accuracy of a perceptual decision. J Vis [Internet]. 2005 May 2;5(5):1. Available from: https://doi.org/10.1167/5.5.1

126. Cardiac Cycle | Anatomy and Physiology II [Internet]. [cited 2023 Jul 12]. Available from: https://courses.lumenlearning.com/suny-ap2/chapter/cardiac-cycle/

127. Sembroski E, Sanghavi DK, Bhardwaj A. Inverse Ratio Ventilation. Mechanical Ventilation: Clinical Applications and Pathophysiology [Internet]. 2023 Apr 6 [cited 2023 Jul 12];256–65. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535395/

128. Mathôt S. Pupillometry: Psychology, physiology, and function. Vol. 1, Journal of Cognition. Ubiquity Press; 2018.

129. Nagai Y, Critchley HD, Featherstone E, Fenwick PBC, Trimble MR, Dolan RJ. Brain activity relating to the contingent negative variation: an fMRI investigation. [cited 2023 Jul 12]; Available from: www.sciencedirect.com.

130. Kloth N, Itier RJ, Schweinberger SR. Combined effects of inversion and feature removal on N170 responses elicited by faces and car fronts. Brain Cogn [Internet]. 2013;81(3):321–8. Available from: https://www.sciencedirect.com/science/article/pii/S0278262613000031