UNIVERSITY OF PADOVA

Department of General Psychology

Department of Developmental Psychology

Master's degree in "Cognitive Neuroscience and Clinical Neuropsychology"

Final dissertation

An electroencephalographic investigation of the impact of eye movements in a change detection task

Supervisor:

Professor **Roberto Dell' Acqua**

Candidate: **Alberto Petrin**

Student ID number: **2035731**

Academic Year: 2022/2023

INDEX

To my beloved family, whose love and support have been my foundation. I hope to be able to return at least a small part of what I have received. To Irene, my pillar and my strength, my happy place, and my motivation over all these years.

To Roberto Dell' Acqua and Mattia Doro, for their mentorship and inspiration to pursue my goals.

Abstract

In studies involving Event-Related Potentials (ERPs), ocular artifacts such as blinks and saccades can compromise the quality of the recorded neural signals. To address this issue, researchers often manually reject epochs (that is a specific time-window extracted from the continuous EEG signal) containing these artifacts. However, this procedure consistently reduces the number of epochs that can be used for extracting ERPs. An alternative solution is to use Independent Component Analysis (ICA), which can preserve more epochs for analysis by removing only the artifact from the EEG recording. However, the reliability of ICA in neurocognitive studies of lateralized ERP components, such as the Sustained Posterior Contralateral Negativity (SPCN) related to visual working memory load, remains unclear, particularly in contexts where subjects are more likely to make saccades during the task. Furthermore, by using ICA, we are assuming that ocular movements do not interact with the neural signal, which has yet to be confirmed. For this reason, in the present experiment, all the participants were asked to perform a change detection task under two conditions: a 'free gaze/saccade' condition, where they were allowed to move their eyes to look at the lateralized stimuli, and a 'fixation' condition, where they were required to maintain the gaze on the center of the monitor. The subjects were also split into two groups, each performing the same experiment but with different stimulus presentation times (100 ms and 500 ms) to investigate whether saccades could differently affect the ERP in these conditions. The SPCN components were then extracted using both the Independent Component Analysis (ICA) correction and epoch-rejection methods. The results revealed that ICA correction is a robust and reliable method for experimental paradigms with a short presentation time of the stimuli (100 ms). By removing only the saccades, the features of the SPCN are preserved, suggesting that with this method we can retain a higher number of epochs for the ERP extraction with the certainty that saccades do not alter the neural signal.

1. INTRODUCTION TO EEG

Researchers of neurocognitive psychology have always been interested in instruments that allow the measurement of neural activity to understand the cognitive and perceptual mechanisms that take place in the brain. Thanks to a series of scientific and engineering achievements started by the Italian scientist Luigi Galvani in 1791 (Galvani, 1791), a technique to directly measure the brain's electrical activity has been developed: electroencephalography.

Electroencephalography consists in the recording and interpretation of the electroencephalogram (EEG), which is typically defined as the electrical activity produced by the firing of neurons in the human brain, and is normally recorded at the brain scalp (Blinowska & Durka, 2006). To acquire high-quality EEG signals, a comprehensive EEG measurement system necessitates several elements: the electrodes (that can be divided into active, reference, and ground electrodes) with conductive medium (gel or saline), amplifiers with filters (to amplify the signal of interest while lowering the voltage from other sources), analog-to-digital (A/D) converter, and recording device (Teplan, 2002).

The number of electrodes used can range from 1 to 256 (high-density electrode arrays can be useful to improve the spatial resolution), but some researchers demonstrated that recording from 32/64 active electrode sites is appropriate for most experiments (Luck, 2014). The standard protocol for electrode placement is called "The International 10–20 System", established by the International Federation in Electroencephalography and Clinical Neurophysiology (see Figure 1). In this protocol, the electrodes are placed at 10% and 20% intervals along lines of both latitude (from the left preauricular point to the right one) and longitude (from the inion to the nasion). The name of each electrode contains two parts, that is one or

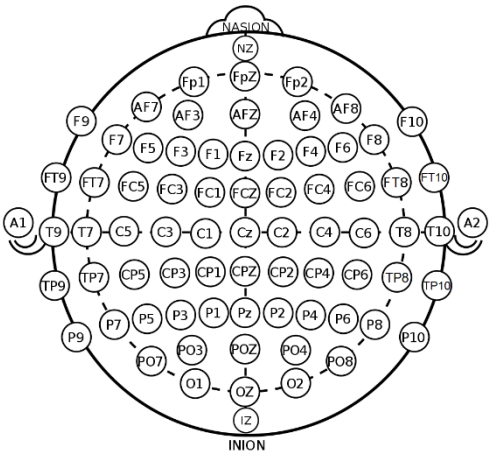


Figure 1: Electrode's positions with an EEG system with 64 active electrodes, following the 'International 10-20 system'.

two letters and a number. The letters stand for the brain region to which the electrode corresponds, with abbreviations such as ‘Fp’ for frontal pole, ‘F’ for frontal, ‘C’ for central, ‘P’ for parietal, ‘O’ for occipital, and ‘T’ for temporal. The numerical component indicates the electrode's distance from the midline: larger numbers indicate a greater distance, with odd numbers positioned in the left hemisphere and even numbers situated in the right hemisphere.

It is demonstrated that the EEG arises from summed synchronized synaptic activities in populations of cortical neurons, with a main contribution from pyramidal cells, oriented perpendicular to the cortical surface. The postsynaptic potentials, which are voltages induced by the binding of neurotransmitters to their receptors on the membrane of the postsynaptic neuron, are considered the primary contributors of EEG activity due to their longer duration (50–200 ms) and the greater potential field compared to the action potentials (Hu & Zhang, 2019).

Another important aspect of EEG is that the voltage fluctuation measured at any electrode on the scalp will be the sum of activities produced by numerous neural sources: this phenomenon is known as spatial smearing of the signal, which determines the poor spatial resolution of EEG (Hu & Zhang, 2019).

Hence, electroencephalography is a technique that provides a direct measure of electrical brain activity generated by thousands of synchronized neurons, with excellent temporal resolution but poor spatial resolution. In addition to the neural signal, EEG also records a lot of noise. All the non-EEG signals (noise) are referred to as artifacts, and they usually have higher amplitudes and distinct morphologies compared to neural signals recorded on the scalp.

Nonetheless, there are a lot of doubts about how to correctly deal with the noise in EEG recordings, and the most effective method to do so has yet to be found. Understanding and distinguishing between the source of noise and the signal of interest is essential for effectively approaching and interpreting the EEG results in the best way possible, and this is what we tried to establish in this study.

1.1. EEG ARTIFACTS

Artifacts in EEG recordings refer to signals that are not generated by the brain and have the potential to mask neural activity, preventing an accurate EEG interpretation. These are commonly divided into non-physiological artifacts and physiological artifacts (Sazgar & Young, 2019).

The most common non-physiological artifacts are: power line interference at 50/60 Hz, impedance fluctuations, cable movements, broken wire contacts, low battery, and near electronic devices interference (see images in Figure 2). These are caused by environmental or system noise, and they can be corrected or prevented with adequate precautions, knowing their systematic appearance (Sazgar & Young, 2019; Teplan, 2002).

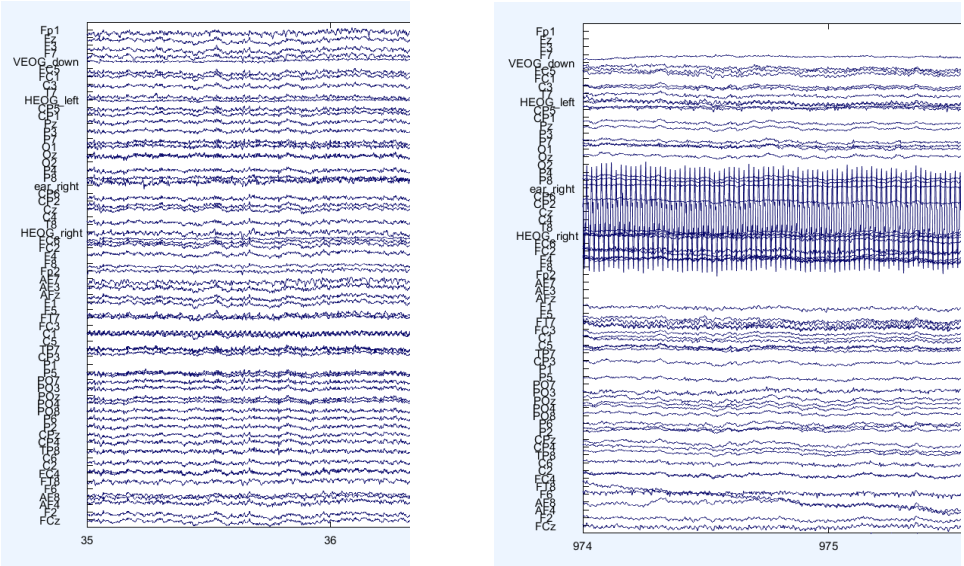


Fig 2: Two examples of non-physiological artifacts: to the left, the power line interference at 50 Hz can be observed in each channel, because of the noisy/non-smooth signal; to the right, channel noise due to impedance fluctuations can be observed.

The physiological artifacts instead include: muscular artifacts, head movements, cardiac artifacts (Figure 3), respiration artifacts, artifacts caused by minor body movements, glossokinetic artifacts (caused by talking and tongue movements), sweating artifacts, and eye movements artifacts such as blinks, horizontal and vertical eye movement (Sazgar & Young, 2019).

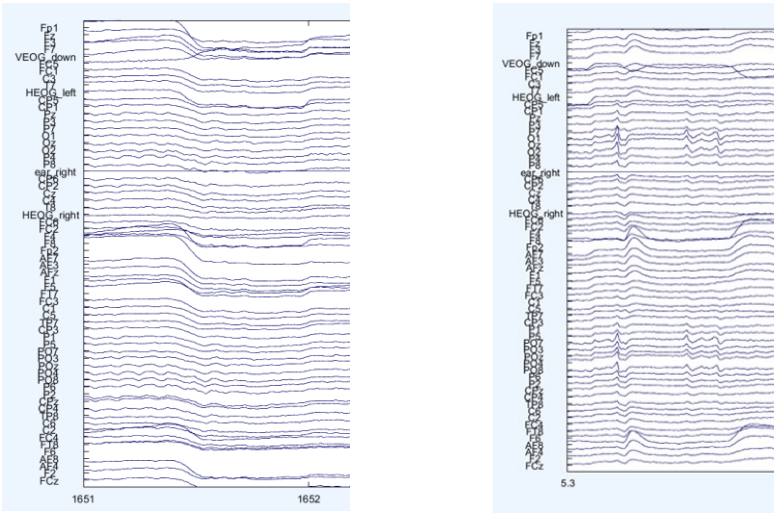


Figure 3: To the left, an example of physiological artifacts due to a head movement. The movement affects all the channels. To the right, an example of a cardiac artifact. It can be observed through the small spikes present especially in the posterior electrodes.

Some artifacts are tiny and steady, like cardiac activity and the power line interference; others are bigger and with a slow frequency resulting in drift in the signal like when the patient is sweating or slowly moving the head; and others, instead, are large and transient like blinks and ocular movements (Hu & Zhang, 2019). Especially in cognitive research, there is one particular kind of artifact that is difficult to avoid and handle, because of its strong link to almost all our cognitive processes, and its highly frequent occurrence: the ocular movements (Figure 4).

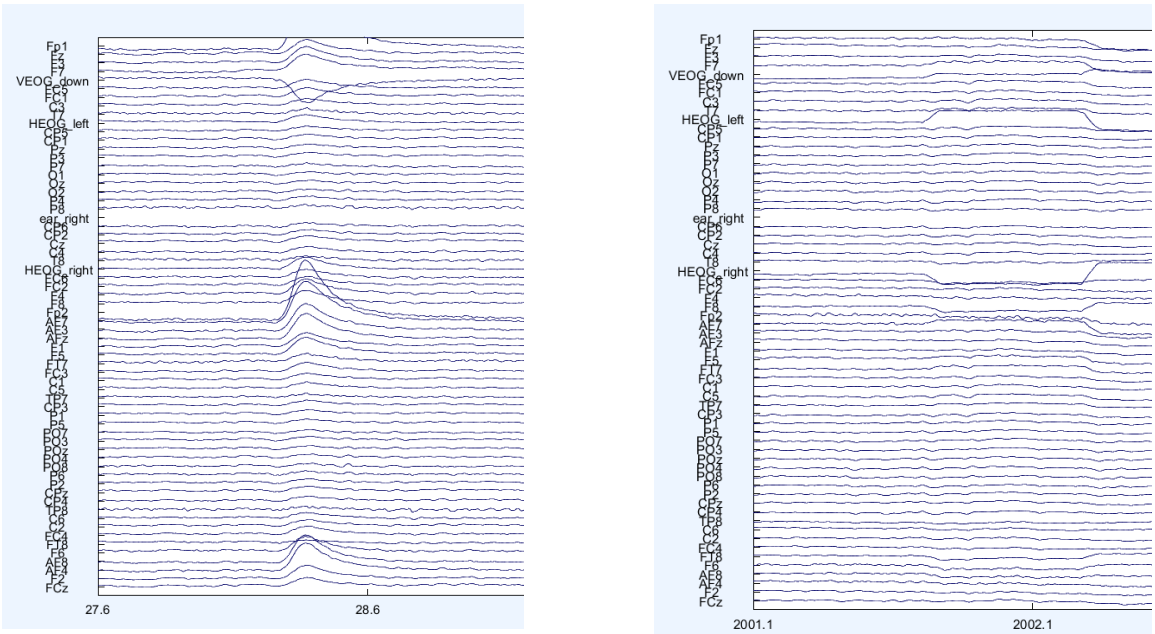


Figure 4: to the left, a blink artifact with a monophasic wave and opposite polarity between the frontal channels and the 'VEOG_down' channel. To the right, a saccade artifact can be observed especially at channels 'HEOG_right' and 'HEOG_left', with opposite polarity.

1.2. OCULAR ARTIFACTS IN COGNITIVE NEUROSCIENCE

The ocular artifacts we observe in EEG can be divided into blinks and saccades.

Blinks are physical cut in our visual continuity, and in normal situations they spread tears across the eyes to eliminate the irritants and foreign objects. In some cases, they can be caused by a potential threat (such as a sudden loud sound or a moving object toward us) like in the startle reflex (Bradley et al., 1999). Blinks can be defined as semi-autonomic rapid eyelid closures because they are partially under our control (for instance, we can change the blink rate and close the eyelid when we want to), but we cannot cease blinking for a long period, as attempting to do so would lead to discomfort. In EEG recordings, the blinks have an amplitude of roughly 100-200 μV (microvolt) while the neural signal typically falls below 100 μV. A blink is characterized by an increase or decrease in amplitude (monophasic waveform) generally within 100 and 500 ms and it is recorded by all the electrodes, but it is particularly evident in the signal from frontal channels like “Fp1” and “Fp2” (Chang et al., 2016). During experiments, EEG is recorded from subjects while they are performing a task in front of a computer, for at least 30 min; thus, it is not possible to ask participants to refrain from blinking because prolonged computer use tends to irritate the eyes, causing eye strain (Sheppard & Wolffsohn, 2018). For this reason, the only way to prevent as much as possible blink contaminations in the EEG is to ask participants to blink only in prespecified moments (for instance after responding or during time windows of non-interest for the study). Nevertheless, this could be considered as an additional task (Drisdelle et al., 2017) and the possible implication of that will be discussed below.

Saccades, due to the visual restrictions imposed by the fovea (1.5 degree diameter), are essential to bringing the selected portion of the retinal image to the area with higher visual acuity (fovea) or to maintain stable gaze on moving objects (Kowler, 1995). The eyes move in a conjunctive way (the movements in each eye have the same amplitude and direction) and the eyeball is relatively light and mobile. As a result, the metabolic costs of very frequent and rapid movements are very low and this suggests that saccadic sampling is a significant aspect of vision (Gilchrist, 2011). These fast eye

movements are typically followed by a fixation which is a period during which the eye remains fixed. Saccades last for about 50 ms, and their average latency can vary from as little as 100 ms to as much as 1000 ms because it is significantly influenced by the nature of the stimuli that must be responded to (Gilchrist, 2011). We can distinguish between 'top-down' and 'bottom-up' factors that determine which area of the space is selected for the next fixation. A 'bottom-up' signal emerges from the visual input and attracts the eyes to the stimulus location, regardless of the task being performed at the time. Conversely, the 'top-down' control allows the eyes to be directed to locations that are task-relevant regardless of their visual salience (Gilchrist, 2011). The movement of the eye blurs the image on the retina, probably functioning as backward masking because it may limit the ability to report visual information from the fixation before the saccade (Gilchrist, 2011). Saccades have an EEG amplitude of 20-80 μV, and they are associated with the deployment of spatial attention toward the object of interest. However, we can allocate attention to objects without making saccades and this ability is called covert attention (Fawcett et al., 2015). This ability allows subjects to maintain fixation on the center of the monitor during cognitive tasks, while still effectively directing their attention to relevant lateralized objects.

It must be noted that eyes are not completely stationary even during the fixation phase of the saccade. There are three types of miniature movements during fixation: tremor, drift, and microsaccades (Carpenter, 1988). All three types of movements play a role in reducing neural adaptation, which prevents the visual picture from fading (Martinez-Conde et al., 2004). Microsaccades can have a small impact on EEG measurements, including both posterior sites (Dimigen et al., 2009) and lateralized ERPs (e.g., SPCN/CDA, see Kang & Woodman, 2014).

As previously mentioned, refraining from blinking and moving the eyes during a neurocognitive experiment could be considered a secondary task that affects the primary cognitive activity of interest (Drisdelle et al., 2017). For example, Verleger in 1991, demonstrated that "the harder some frequently blinking subjects try to refrain from blinking, the smaller might become their P3 amplitudes"

(Verleger, 1991). Therefore, the influence of saccadic movements on ERP latency or amplitude remains a topic of ongoing investigation and is not yet fully understood.

This is a significant concern considering that there are ERP components whose amplitude is modulated by the manipulation of some experimental variables (e.g., SPCN component, see subchapter 2.2): if the amplitude or latency of an ERP is modified not only by the independent variable of interest but also by other covariates of non-interest, there is a risk of misinterpreting the results.

To gain a better understanding of the problems and the solutions adopted to avoid the impact of eye movements and the relative artifacts, it is crucial to have a clear idea of what an ERP is.

2. THE ERPs

"Event-Related Potentials" (ERPs) are waveforms of electrical activity generated by the brain after the occurrence of an event, which could be the presentation of a particular stimulus or the execution of a response (Bradley & Keil, 2012; Hu & Zhang, 2019). Usually, ERP components are named based on their polarity (positive or negative) and their latency (the time after the event at which they occur). For example, the N200 component is a negative ('N') deflection in the EEG that occurs about 200 milliseconds after the presentation of a stimulus. However, other components such as the SPCN, N2pc, or MMN, are named also considering the polarity, the spatial localization of the activity, or some other features of the waveform (e.g., 'Sustained Posterior Contralateral Negativity', 'Mismatch Negativity').

The ERP components are thought to reflect specific sensory-cognitive processes such as attention, memory, and decision-making. ERP technique is useful because EEG data have a low signal-to-noise ratio (SNR), and it cannot be utilized in its raw form to measure most of the neural processes related to a specific event or task (Hu & Zhang, 2019). However, by time-locking the neural activity recorded in each trial of the same condition and averaging these for each channel, random noises tend to cancel each other out. Ideally, this leaves only the signal of interest generated by the event: the ERP.

According to the sensory stimuli that are presented we obtain different Event-related potentials. For instance, the presentation of an auditory stimulus evokes potentials generated in the cochlea, known as Auditory Evoked Potentials (see Figure 5). Within a few milliseconds after a sudden sound onset, the first set of auditory responses appears, reflecting the flow of information from the cochlea through the brainstem and into the thalamus. These are called brainstem auditory evoked potentials (BAEPs) and even though they exhibit very low voltage (approximately 0.5 μV in amplitude), they can be used to evaluate the integrity of the auditory pathways (Hu & Zhang, 2019). These waves are followed by the mid-latency responses (MLRs), which encompass responses occurring between 10 and 50 ms after stimulus, and subsequently by other long-latency responses (occurring between 50 and 200 ms).

These long-latency responses include, in sequence: the P50 (also known as P1), N100 (known as N1), and P160/P2 components (Luck, 2014). With auditory stimuli, we can also observe the "Mismatch Negativity" component (MMN) which is thought to indicate an automatic response to a stimulus that differs from the preceding ones. MMN seems to be modulated by attention even though there are studies suggesting that it may arise from an attention-independent process (Sussman, 2007).

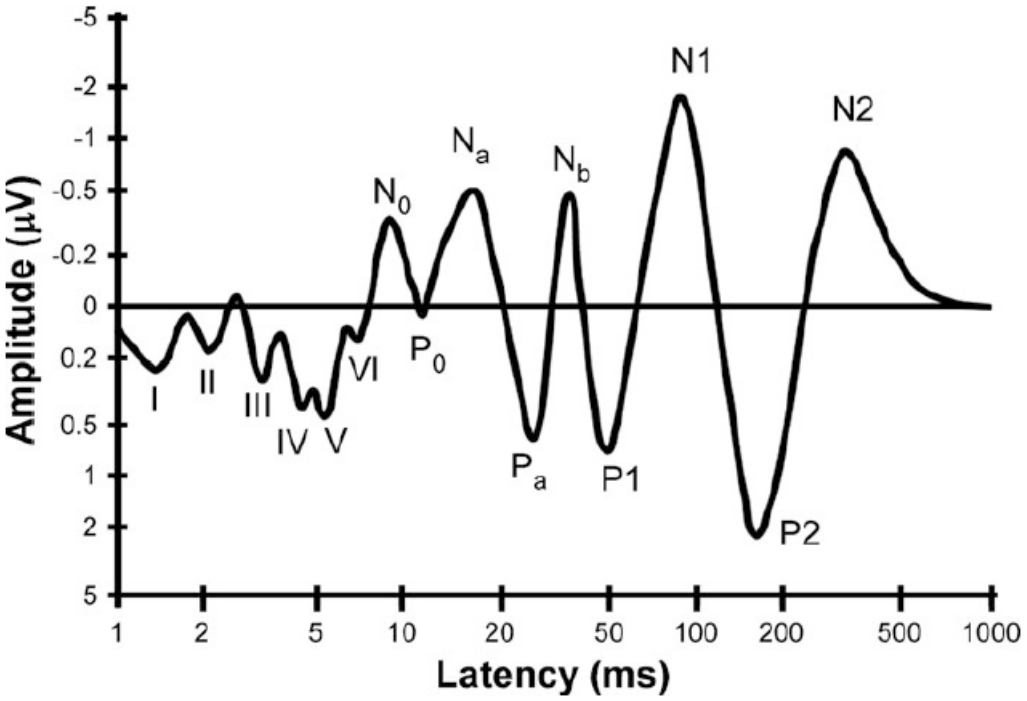


Figure 5: Typical sequence of auditory evoked components. The waveform elicited by an auditory stimulus is represented over a period of time to show the auditory brainstem responses (waves I– VI), the mid-latency responses, and the long-latency responses.

The presentation of visual stimuli, instead, generates 'Visual Evoked Potentials' (Figure 6). The first major visual ERP component is the P1 wave, peaking between 100 and 130 ms at lateral occipital electrode sites after 60–90 ms post-stimulus presentation. It is thought to be generated in the extrastriate visual areas (Di Russo et al., 2002). Following the P1 wave there is the visual N1 wave, which is highly refractory, meaning that the response to the second stimulus is significantly diminished if the first stimulus is presented at the same location after a very short delay. The earliest N1 subcomponent peaks at 100–150 ms post-stimulus at anterior electrode sites. These components,

together with the N2 wave, are influenced by or reflect spatial attention. Following, there are also other components and subcomponents, such as: the P300, related to working memory and attention; N400 and P600, which are associated with semantic and syntactic language control respectively (Hu & Zhang, 2019).

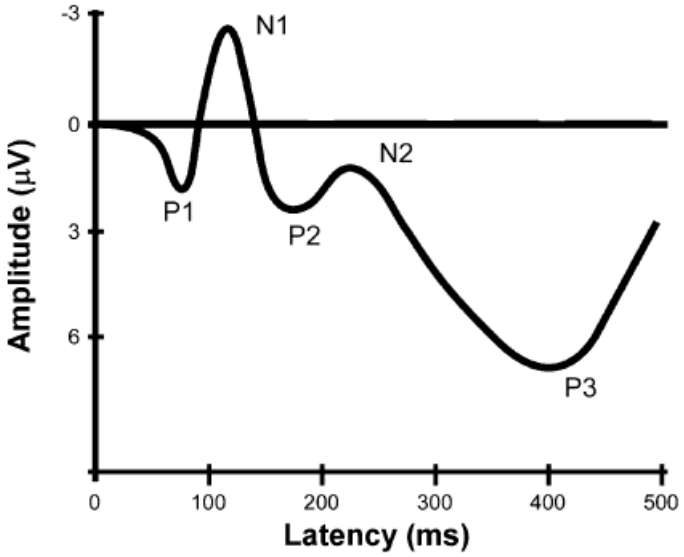


Figure 6: Typical visual evoked components, including P1, N1, P2, N2, and P3 components.

These are only a few examples of ERPs that we can observe, but there are many more in the literature. For instance, other ERPs are observed in response to: electrical stimulation of fibers in the peripheral nerve (for 'Somatosensory evoked potentials' see Cruccu et al., 2008), infrared stimulation of nociceptors in the superficial layers of the skin (for 'Laser evoked potentials' see Bromm & Treede, 1984), stimulus anticipation (for "Readiness Potential" see Schurger et al., 2021), and so on. ERPs can represent both higher and lower cognitive functions.

2.1. LATERALIZED ERPs

Another way to categorize ERP components is by considering the asymmetric pattern of activity, like for the “lateralized ERPs”. The word “lateralized” means that the presentation of the stimulus of interest, which is often visual, is presented only on a specific side of the monitor (either the right or left hemifield), leading to a difference in the electrical activity of the contralateral hemisphere to the stimulus compared the ipsilateral hemisphere. Specifically, when we subtract the ipsilateral ERP of one channel from the contralateral ERP of the oppositive channel (e.g., PO7 and PO8), we typically observe an increase in the negativity of the waveform amplitude at scalp recording sites contralateral to the visual hemifield occupied by the target stimulus (Roy & Faubert, 2023; Vogel & Machizawa, 2004). These lateralized ERPs seem to be associated with the allocation of resources for specific cognitive processes (e.g., spatial attention or visual working memory) toward relevant objects on the side of the screen. For instance, one well-known lateralized component is the ‘N2pc’. The N2pc is typically observed between 180–300 ms following stimulus onset at posterior electrodes contralateral to an attended visual field, and it is strongly related to the deployment of visual-spatial attention toward lateral targets (Eimer, 1996; Luck & Hillyard, 1994).

2.2. SPCN/CDA

Another relevant lateralized ERP component is the so-called ‘SPCN’ or ‘CDA’. ‘CDA’ stands for “Contralateral delay activity” (Vogel & Machizawa, 2004), and ‘SPCN’ for “Sustained Posterior Contralateral Negativity” (Jolicœur et al., 2006): different names that represent the same ERP component. From now on, we will refer to this component as SPCN. The common way to elicit a SPCN component is by presenting lateralized objects accompanied by a visual cue displayed before the stimuli to indicate which side of the screen is task-relevant. The so-called ‘contralateral control method’ allows isolating the visual working memory (VWM) specific activity by subtracting the ipsilateral from the contralateral activity elicited by the to-be-remembered stimuli (Figure 7). In this

way, common low-level sensory and perceptual processes are removed, isolating the VWM-specific ERP activity. The SPCN is observed at posterior electrodes, it begins roughly 300–400 ms post-stimulus onset and appears to be dissociable from the N2pc. A particular aspect of this component is its sustained activity during the retention interval before the recall phase (Vogel & Machizawa, 2004). From the literature emerges that SPCN amplitude correlates with the number of to-be-remembered stimuli, reaching a plateau around the estimated capacity of 3/4 elements (Luria et al., 2016). SPCN amplitude also increases for complex stimuli (Luria et al., 2010) and faces (Sessa et al., 2011), supporting the idea that complex objects require more VWM resources. This component has been associated with the parietal cortex, in particular the intraparietal sulcus (Becke et al., 2015). Hence, SPCN seems to be associated with the maintenance of information in visual short-term memory and it is sensitive to several VWM features, such as capacity, precision, and distractor filtering (Luria et al., 2016).

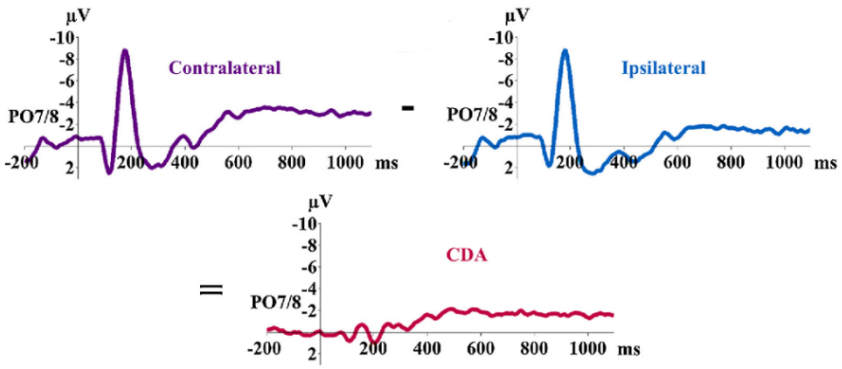


Figure 7: The CDA is time-locked to the onset of the memory array, and the activity is measured throughout the retention interval.

2.3. RELEVANCE OF SPCN

Studies investigating how saccadic eye movement activity can affect posterior lateralized ERP components, already exist in literature (Drisdelle et al., 2017). However, the focus of Drisdelle's study was primarily on the N2pc component that is relatively transient compared to the SPCN (which lasts for the whole duration of the retention interval). Hence, the probability of saccadic eye movements

and blinks occurring during the time window of interest is higher for the SPCN, and if not properly handled, these artifacts can significantly distort the waveform.

Besides, since SPCN is sensitive to different levels of memory load, and we currently have no information regarding how ocular artifacts influence SPCN amplitude at different set sizes of the memory array, this is an extremely important aspect to investigate.

Given that the SPCN component is measured at posterior electrodes (with the greatest amplitude observed at PO7 and PO8 electrodes), it could be that the influence of ocular artifacts' activity is too small to affect the waveform. Anyway, volume conduction (which refers to the transfer of electrical potentials to a site a distance away from the generator) of large ocular artifacts could still be sufficiently large to contaminate EEG data at the posterior sites. Indeed, the amplitude of a blink and saccade is in the order of hundreds or dozens of μV while the amplitude of this component is usually not larger than 5 μV (Drisdelle et al., 2017; Hu et al., 2010).

The SPCN is a very informative component to investigate VWM thanks to its sensitivity to different features and quantities of stimuli. It is, therefore, crucial to establish the best practice for extracting it, and assessing its reliability in the presence of saccadic movements and blinks. All the ERPs are extracted following a similar procedure that is based on an important principle: the more trials available for averaging, the higher the SNR and the better the ERP component. Therefore, it is imperative to enhance the SNR by removing the artifacts, to obtain accurate and clear ERPs.

3. PROCEDURE AND CRITICS

The preprocessing of the raw EEG data is a delicate and crucial step in the ERP extraction because little mistakes or inadequate techniques in the procedure could ruin completely the Event-related potential. The most important steps in the preprocessing pipeline are (Hu & Zhang, 2019):

1. **Import data**: usually, a Matlab toolbox such as "EEGLAB" or "BRAINSTORM" is needed to upload the digital version of the EEG recordings, but it can also be used some R packages (for example, "eeguana" and "eegUtils", see Nicenboim, 2020).
2. **Import channel locations**: this step is needed to load the channel location file describing the channel positions on the head. This file can be created specifically for the study or found inside the "EEGLAB" folder, since all the EEG montages follow "The International 10–20 System".
3. **Filter data**: Filtering is often necessary due to several reasons, such as the presence of 50 Hz or 60 Hz line noise, high and low frequency noises. As a result, digital filters that employ the "Fourier transformation" can be applied to the raw EEG recordings and they can significantly enhance the quality of EEG signals, otherwise heavily contaminated by environmental artifacts. Usually, it is used a "bandpass filter" to maintain only the frequencies between 0.01 Hz and 30/45 Hz (according to the kind of neural signal of interest), and a "notch filter" or "band-stop filter" to remove the 50/60 Hz frequency caused by the power line.
4. **Re-reference data**: basically, EEG consists in the detection of current potential differences between one (or more) active electrode (A) and an online reference (R). The presence of the ground electrode (G) is fundamental to obtaining an output ('channel 1'= (A - G) - (R - G) = A - R). It is also possible to re-reference the data offline. In this case, we use as online reference only one lateralized electrode, like the right/left mastoid or right/left earlobe, and the average of the two earlobes/mastoids is commonly used as re-reference to avoid lateralization bias in the data (Hu & Zhang, 2019).

5. **Check and interpolate bad channels**: it is possible that some electrodes may not properly record the signal due to factors such as sweat, a bad montage, or technical issues. During this step, electrodes that exceed a certain threshold (e.g., three standard deviations from the mean) are interpolated with the signal of the neighbor electrodes.
6. **Segment the EEG data and extract epochs**: EEG data are time-locked to the onset of specific events of interest and then are segmented, to identify changes in EEG activity, at the onset of sensory stimulation. In this phase, the selection of the event onset (the time point "0") and the time windows (how much time before and after the event onset to consider) are essential to segment the continuous EEG recording. These segments are called epochs.
    - **Baseline correction**: removing the mean of baseline values from each epoch is necessary since the electrical potential at baseline varies across data epochs. Baseline correction serves as a potential alternative to strong high-pass filtering and the baseline period to consider should be that before the external event occurs.
7. ***Epoch rejection***: EEG epochs that are greatly contaminated by artifacts (e.g., eye blinks and movements) are marked as bad epochs and then rejected. This phase is important to increase the SNR for the final step, and it could be done manually (selecting the "bad epochs" by visually analyzing them), or automatically, considering the peak-to-peak amplitude in the EEG data and selecting only the epochs that exceed a pre-defined threshold. This threshold cutoff is applied to the electrooculogram (EOG) based on the characteristics of the ocular activity (Drisdelle et al., 2017). For what concerns blink, there is a monophasic deflection often exceeding 100 μV and with an opposite polarity for electrodes placed above (e.g., Fp1, Fp2) and below (vertical EOG, or VEOG) the eye. Regarding the saccades, there is a large increase in positivity at the electrodes on the same side as the foveated visual field relative to electrodes on the other side (e.g., if we look right, a positive-going deflection is observed in the right horizontal EOG or HEOG). Both automatic and manual methods have flaws: the manual one is time-consuming, and the results can vary a lot according to who is doing it; while the

automatic method could use criteria not appropriate for all subjects, raising both type I and II errors. In any case, the automatic procedure is the most used.

8. **Average of epochs to obtain the ERP**: as last step, all the epochs belonging to the same condition are averaged, and if the noise/artifacts correction has been done correctly, we obtain event-related potential for each condition in each channel. Additionally, a "Grand Average" can be done, firstly creating an average ERP for each subject and then, averaging them together.

However, 'step 7', involving epoch rejection, can be a particularly problematic step in EEG data analysis. Indeed, as we have already mentioned, having more epochs per condition is generally desirable because it leads to a reduction in noise and a higher SNR in the ERP, which in turn can improve the reliability of the results.

Although this methodology is necessary to ensure data quality, researchers often reject as much as 50% or more of the data for any particular subject (Drisdelle et al., 2017). Therefore, the experimental design must include many more trials than would be necessary as well as several additional subjects to compensate for expected attrition. Even though this strategy is often used, making the task longer takes a lot of time, and leads to other problems concerning fatigue, a decrease in concentration, and participants' discomfort (Hu & Zhang, 2019). Indeed, considering the EEG montage, the experiment may last more than one hour and a half in some cases. Thus, in the effort to add more trials to reduce the noise, there is the risk of inadvertently increasing noise. Another problem with the epoch-rejection method is that by excluding the trials with ocular artifacts, we are assuming that these artifacts do not significantly interact with the perceptual and cognitive effects being observed in EEG data (Drisdelle et al., 2017). However, this assumption may not always be true.

Furthermore, despite carefully developed EEG experimental paradigms, epoch rejection might lead to unbalanced designs due to unequal rejection rates between conditions. When different conditions have varying numbers of rejected trials, it can lead to differences in variances between conditions, creating a homoscedasticity problem, with difficulties in comparing them statistically. This aspect is

particularly relevant in studies of visual-spatial attention and memory, especially when participants are presented with stimuli in the visual periphery. In such experiments, participants may naturally saccade to task-relevant stimuli, even when instructed to maintain fixation on the center of the screen (Drisdelle et al., 2017), leading to unequal rejection of epoch among subjects and conditions.

Alternative or integrative techniques must be considered to deal with these issues.

3.1. ALTERNATIVE APPROACHES

The best way to resolve or attenuate the problems of the epoch rejection technique (e.g., too few trials to extract good ERPs and inhomogeneous variances between conditions), would be a method that manages ocular artifact by isolating and removing the eye movement-related activity, while preserving the underlying neural sources (Drisdelle et al., 2017).

Several well-known approaches can be employed for this purpose, including: linear regression (Schlögl et al., 2007), dipole modeling (Berg & Scherg, 1991), Principal Component Analysis (PCA), and Independent Component Analysis (ICA).

The linear regression is a statistical technique that can be used to model and remove ocular artifacts. By regressing out of the EEG data the ocular artifacts, it is possible to estimate a propagation factor between the EOG (electrooculogram) channels and each EEG scalp channel and subtract the proportion of activity corresponding to the EOG activity from the ERPs at each scalp site. This is possible because researchers have demonstrated that both eye blinks and horizontal saccades propagate through the head via volume conduction in a linear way (Plöchl et al., 2012), though differently on the scalp. With this method, EOG artifacts can be reduced by 80% and it can be used with any number of electrodes (Schlögl et al., 2007). The problem with this technique is that EOG channels contain not only ocular activity; indeed, the risk is to also remove neural signals of interest

from the EEG data. The neural signal loss appears to range from about 5% for frontal EEG channels to less than 1% for occipital channels (Schlögl et al., 2007).

Another approach to the ocular artifacts' correction is dipole modeling. This model, originally proposed by Berg and Scherg, is based on the understanding that the retina is electrically charged (positive at the front and negative at the back). For this reason, it assumes that "the electromagnetic field can be represented at distances corresponding to those of sensors placed on the head and around the eyes by an equivalent current source dipole located somewhere in each eye and oriented approximately in the direction of gaze" (Berg & Scherg, 1991). Effects of the dipole are undetected if the eyes are still, but when the eyes move, the effects of changes in position or orientation of the ocular dipole can be detected. Basically, this procedure models spatiotemporal dipoles by assuming a priori the number of dipoles related to ocular movements (e.g., blinks, saccades). The problem in this case arises from the "a priori assumptions": if they are incorrect, this may lead to subsequent inaccuracies in the source's locations and the contributions of EOG to EEG (Drisdelle et al., 2017).

The PCA is another option. Principal Component Analysis finds the components in the data that explain the greatest variance (the order in which they are presented is due to the proportion of variance explained by each component) and that are orthogonal to each other. Therefore, PCA aims to extract the smallest set of variables with less redundancy, where redundancy is defined by correlations between data elements (Bugli & Lambert, 2007). This procedure is a dimensionality reduction technique that allows the simplification of the data in multivariate statistics. This means that it is not the best tool to find distinct components in the data, but rather to identify underlying variables that are uncorrelated with each other. In fact, if some components are uncorrelated, it does not mean that they are also independent: as can be seen in Figure 8, knowing the values of x provides information about y, even though x and

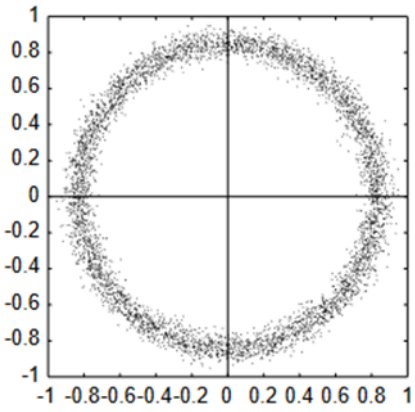


Figure 8: Uncorrelated but dependent variables.

y are uncorrelated. This is problematic since PCA can extract a large set of different but statistically equivalent uncorrelated/orthogonal components, failing to accurately identify the original individual components (Stone, 2002). Additionally, since PCA generates a series of orthogonal base vectors where each vector accounts for as much variance as possible (Bugli & Lambert, 2007), the first vector will result significantly larger in magnitude than all the subsequent ones. When the SNR is low, important information in these subsequent vectors can get lost. Moreover, it has been seen that PCA does not completely separate artifacts from neural activity, especially when they have similar amplitudes (Lagerlund et al., 1997).

Thus, the last technique to consider is the Independent Component Analysis, and as we are going to see, it seems to be the more appropriate to use for artifact correction in EEG data.

3.2. ICA

ICA was originally developed to deal with the cocktail party problem, that is the attempt to isolate a pertinent conversation from the noise of other conversations in a cocktail party (Hyvärinen & Oja, 2000). Consider a party where many people are talking at the same time. If there is a microphone, then its output is a mixture of voices (X). Given such a mixture, ICA identifies those individual signal components (Ũ) of the mixture that are unrelated. Since the only unrelated components within the signal mixture are the voices of different people, this is what ICA identifies (Stone, 2002). It is worth stressing that ICA requires more than one simultaneously recorded mixture in order to find the components and it does not include any information specific to speech signals (Stone, 2002).

The effectiveness of ICA is based on the assumption that source signals are not only uncorrelated but are also ‘statistically independent’, which means that the value of one variable provides absolutely no information about the value of the other (Hyvärinen & Oja, 2000). ICA also relies on the assumption that signal mixtures, which are combinations of independent source signals, must have

non-Gaussian distributions (Hyvärinen & Oja, 2000). Hence, ICA is a reliable instrument to deal with EEG signals, since EEG data might not always meet the normality assumption (Sugimoto et al., 1978), and consist in a mixture of activity produced from different physical processes in various brain regions (neural and nonneural processes). Must be said that ICA is a technique of dimensionality reduction, but differently from PCA, it first separates the independent components and then removes those of no interest from the data. It can be considered a generative model, which means that it describes how the observed data are generated by a process of mixing the components (Hyvärinen & Oja, 2000).

Through the ICA formula, we can observe that the final signal "X" (with EEG data, it consists in a matrix 'channels*time') can be expressed as a linear combination of independent components "U" (a matrix 'components*time') and an unknown mixing matrix (A):

$$X = A*U$$

Since we are interested in extracting the independent components that compose the final signal, ICA defines a separating/demixing matrix "W" (the inverse of the estimated mixing matrix A) to recover an estimation of the components ($\tilde{U}$), multiplying it with the mixture signal (Hyvärinen & Oja, 2000):

$$\tilde{U} = W*X$$

Thus, while in the original EEG data each row of the recording data matrix represents the time course of activity of the channel (channel = A – R), after ICA decomposition each row of the transformed data matrix represents the time course of the activity of one independent component (IC), spatially filtered from the channel data. These outputs reveal information about the temporal and spatial characteristics of the ICs. To remove artifacts embedded in EEG recordings, the computed ICs are first classified as either artefactual or neural components (Zou et al., 2016). If detected and flagged as artifact-related ICs, they can be subtracted from the recorded data putting to '0' the column of the mixing matrix (A) associated with the flagged artifact, and the remaining data can be remixed. In ICA

correction, artefactual ICs can be identified thanks to their characteristic shapes (topographies, time courses, and frequency spectra) and can often be classified automatically.

It must be also said that ICA is a special case of "Blind Source Separation" (BSS). "Source" means an original signal (i.e. independent component), and "Blind" indicates that we have little to no knowledge of the mixing matrix (Hyvärinen & Oja, 2000). This blind source separation has some drawbacks: it may result in the possibility of variation in the separation of these underlying sources, and the variability in the source related to the artifact may reflect the inclusion of elements of non-artifactual sources. That is, when the EEG signals are reconstructed without the artifact component, there is the potential that portions of the neural signal may have unknowingly been removed (Pontifex et al., 2017). Besides, with ICA we cannot determine the variances as well as the order of the independent components (Hyvärinen & Oja, 2000).

3.3. ICA ARTIFACTS

As already mentioned, artifact components generally can be identified according to the topographies of the electrodes on the scalp, the across-trial temporal distributions (consisting in a matrix 'trial*time' that shows where in time and how often in the trials the artifactual components appear), and frequency distributions of the components (Hu & Zhang, 2019).

Abnormal topographies can appear as (1) power concentrated just in the frontal lobe in topography (ocular artifacts); (2) discontinued topography (noise artifacts); and (3) topography constrained within a single electrode (electrode artifacts). Abnormal across-trial temporal distributions can appear as (1) inconsistent between epochs (without clear peaks in average waveforms); (2) periodic waveform (power line interference); and (3) noisy patterns (like the Gaussian noise). In EEGLAB, the artifact components are already flagged and labeled by ICA, while the rejection of the artifacts can be done automatically or manually. If all labs were to use the same ICA algorithm and function

to automatically reject labeled ICs, it would be the most optimal approach. That is because it would reduce the variability that experimenters could introduce into the data by removing different components, thereby increasing the replicability of the results. However, since ICA is not perfect in labeling components, there is the risk of raising the amount of type I (rejecting an artifact that it is not) and type II (not rejecting an artifact when you should) errors. Because of this, manual rejection is still used, and it may even be more appropriate if the experimenters have enough experience. Nonetheless, a good practice is to follow the instructions above to decide whether a component is or not an artifact, and if we have some doubts about an IC, we shouldn't remove it. Besides, it is advised to not remove more than 2-4 components when we have good quality EEG data (Hu & Zhang, 2019).

Notably, the application of ICA seems to be particularly useful in removing blinks and other oculomotor artifacts (Plöchl et al., 2012). They tend to have an anterior distribution, and their time courses are largely flat with occasional very high-amplitude spikes indicating artifacts of the eye muscles (Figure 9). Using ICA to correct artifacts is generally considered the best (Hyvärinen & Oja, 2000), especially in the case where we also want to investigate whether the cognitive processes change if we move the eyes.

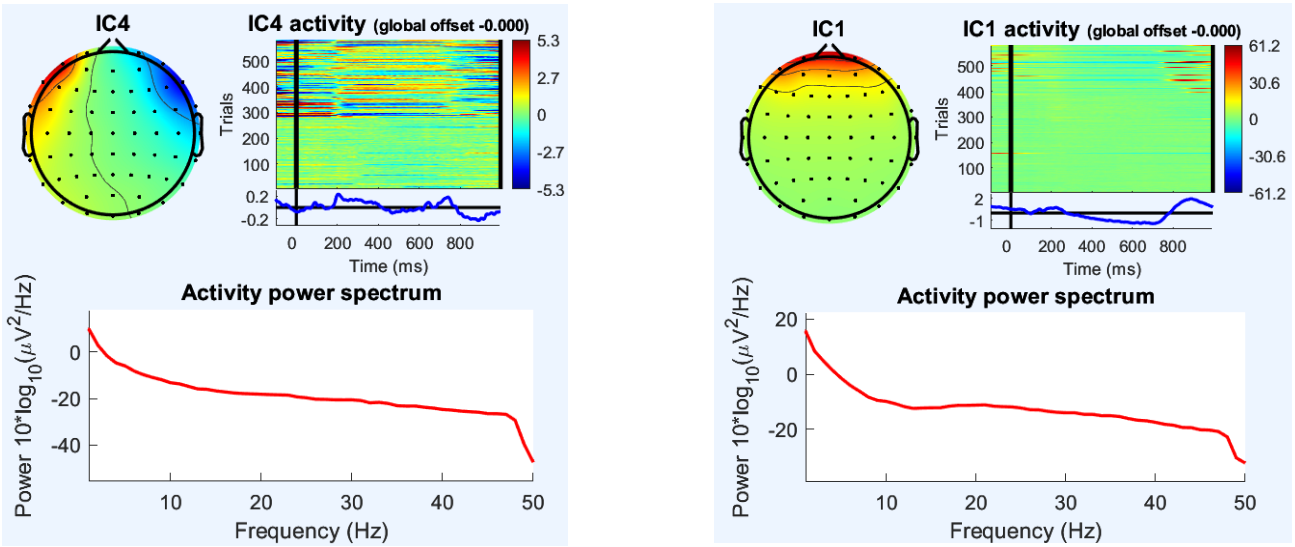


Figure 9: to the left, a saccades component found through ICA. We can appreciate the typical topography and the abnormal across-trial temporal distribution, which represents a higher number of saccades during the saccade block of the task. To the right, a blink component found through ICA, with its typical topography and across-trial temporal distributions inconsistent between epochs.

3.4. ICA IN ERP PIPELINE

Considering the pipeline proposed in ‘paragraph 3’, if we want to reduce the number of rejected epochs, we should perform ICA before the epoch rejection step:

1. Import data;
2. Import channel locations;
3. Filter data;
4. Re-reference data;
5. Check and interpolate bad channels;
6. Segment the EEG data and extract epochs;
    - Baseline correction;
7. ***ICA (artifacts correction)*;**
8. Epoch rejection;
9. Average of epochs to obtain the ERP.

In this way, ICA corrects the artifacts before rejecting epochs that exceed the threshold, increasing the number of trials retained for the extraction of our ERP.

Besides, some studies have shown that according to the algorithm used to compute ICA, some variability may be introduced into the data due to the uncertainty of the separation of the underlying sources/components (Pontifex et al., 2017). The matrix ‘W’ is estimated using an iterative algorithm that maximizes the independence of the estimated sources. The most reliable and used algorithms are called: ‘FastICA’, ‘SOBI’, and ‘infomax’. Pontifex demonstrated in a paper of 2017, that the ‘infomax’ algorithm produces the greatest reduction in the eyeblink artifact, even though it is computationally slower than the others. Furthermore, uncertainty in the ‘infomax’ algorithm seems to produce a small amount of variability in potential solutions compared to the other two algorithms. Anyway, the mechanism underlying the superior performance of this algorithm is unclear. This aspect is extremely important for the goal of our study because, to understand whether ICA is a reliable

instrument and whether moving the eyes influences the underlying cognitive processes, it is essential to ensure that differences do not depend on the ICA algorithm used.

In summary, we have identified the primary challenges associated with ERP studies and have proposed a potential solution to address these issues. Even though several experiments have already validated ICA as an appropriate procedure to remove artifacts for non-lateralized ERPs (Mennes et al., 2010), very few studies tried to examine the effect of saccadic eye movement activity on posterior lateralized ERP components like the SPCN component (Drisdelle et al., 2017).

4. RESEARCH QUESTIONS

The experiment that we performed has been inspired by the paper of Drisdelle and colleagues of 2017, and we aim to assess the effectiveness of ICA for saccade and blink removal on a lateralized posterior ERP (SPCN component). This was accomplished by having the participants complete a "change detection task" (CDT) for visual working memory that was divided into two blocks: (1) one with traditional instructions (fixation condition), where subjects maintain fixation during experimental trials; and (2) another with heavily saccade-contaminated data (free gaze/saccade condition), where subjects are instructed to saccade toward the target. Then, we analyzed the data with both methods (ICA and epoch-rejection method) to determine if ICA somehow distorted the signals, especially when saccades were present. The effects of saccades on SPCN components were evaluated investigating whether the differences in amplitude at different memory set sizes were coherent with the findings present in literature and whether the method of analysis interacted with those. This experiment has a within-subjects design, but we also added a between-subjects variable, that is the presentation time of the memory array (500 vs 100 ms) because the traditional condition with presentation time at 100 ms could be suboptimal for letting subjects saccade toward the target (given that the latency of a saccade is higher than 100 ms, see Gilchrist, 2011). In this way, we also assessed possible differences in SPCN amplitude at different presentation times for each experimental block, between the two methods of analysis.

To sum up, we are going to investigate these main questions:

1) Does ICA produce reliable and solid results compared to using only the epoch-rejection method when extracting SPCN components?
2) Do saccades influence the SPCN waveform? If so, in which conditions?
3) Does a more ecological setup (longer presentation time and possibility to saccade) improve the SPCN components if properly handled with ICA?

5. METHOD

5.1. Subjects

Forty-four students at the University of Padova (30 women, and 14 men) took part in the present experiment after giving formal consent. All participants reported normal or corrected-to-normal vision and had no history of neurological and/or psychiatric disorders. Specifically, twenty-three subjects performed the experiment with the memory array presented for 100 ms, and twenty-one performed the experiment with the memory array presented for 500 ms. Eight subjects were excluded from the analysis (five from the first group and three from the second group): one participant did not reach an accuracy of 60% of correct responses in the change detection task; four participants were removed because the trials remaining after artifact rejection were less than 50%; and other three subjects were rejected because the ERPs presented contaminations of alpha waves and muscular artifacts that disrupted the waveforms. Thus, thirty-six subjects were kept: eighteen for the first group (age: $M = 23.6$, $SD = 2.27$; 6 males; 1 left-handed) and eighteen for the second group (age $M = 23.2$, $SD = 2.03$; 6 males; 1 left-handed). The experimental protocol was vetted by the local Ethical Committee.

5.2. Stimuli and Task

The stimuli of the change detection task were generated with E-Prime 2 software (Psychology Software Tools Inc.) and displayed on the black (RGB: 0, 0, 0) background of a 24” CRT monitor with a refresh rate of 60 Hz at a distance of about 65 cm. An example of the stimuli and a schematic illustration of the sequence of events in the single-probe change detection task (CDT) is reported in Figure 10.

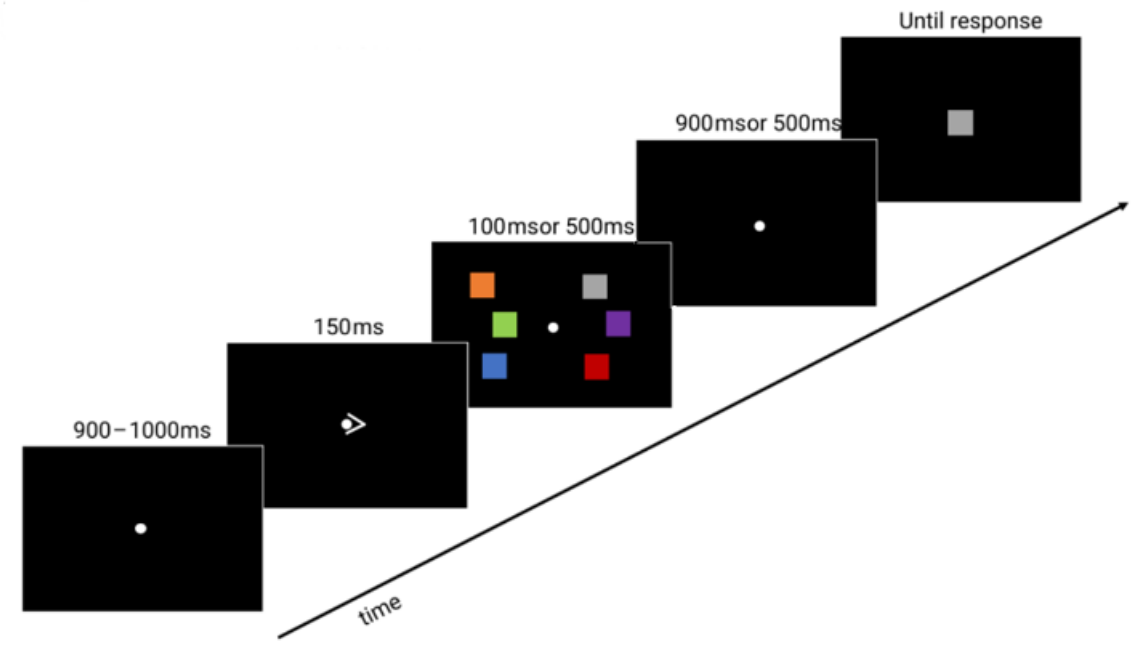


Figure 10: Schematic illustration of the sequence of events in one trial of the single probe change detection task.

Each trial started when the participants pressed the spacebar of the keyboard positioned in front of them. Upon spacebar press, a white (RGB: 255, 255, 255) dot, subtending .8 × .8° of visual angle, was displayed at the center of the screen for a 900–1000 ms interval, randomly jittered in steps of 20 ms. A white (RGB: 255, 255, 255) arrow appeared around the dot for a 150 ms interval, pointing toward the side participants had to attend. Two memory arrays composed of two, three, or five equiluminant colored squares were then displayed to the left and right of the fixation dot for 100 or 500 ms (according to the group the subject was assigned before the experiment started). The cued side could be, with equal probability, the right or left hemifield. Participants had to attend and memorize only the memory array on the cued side. Each square subtended 1° × 1° and the colors were randomly chosen among blue (RGB: 0, 0, 255), brown (RGB: 157, 0, 23), orange (RGB: 255, 128, 0), purple (RGB: 128, 0, 255), dark green (RGB: 30, 140, 60), cyan (RGB: 0, 255, 255), pink (RGB: 255, 174, 201), magenta (RGB: 255, 0, 255), yellow (RGB: 255, 255, 0), red (RGB: 255, 0, 0), and light green (RGB: 0, 255, 0). After a 900 or 500 ms blank retention interval (according to the group the subject was assigned before the experiment started) in which only a white dot at the center

of the screen was present, a test-colored square (single probe) was displayed representing, with equal probability, a color present or absent in the memory array that subjects had memorized. The color of the probe was selected randomly between the colors present in the previous memory array or between the colors absent from the memory array. The probe remained in view until participants pressed one of two keys (i.e., the keys 'A' or 'L' of the computer keyboard, counterbalanced across participants) to indicate whether it was present or absent from the memorized memory array. The experiment was composed of 696 trials divided into two experimental blocks (348 each). For half of the experiment (i.e., one block), subjects were instructed to maintain fixation on the dot at the center of the screen (fixation block). For the other half of the experiment, subjects were instructed to direct their gaze toward target items, that is the memory array in the cued side (saccade block). The order of experimental blocks was counterbalanced across subjects. Every 29 trials there was a pause whose duration was decided by the subjects. Usually, the pause between the two experimental blocks was the longest. Before each experimental block, there were 18 trials of practice. On average, subjects took 1 hour and 15 minutes to complete the task.

5.3. Electroencephalography

The EEG was recorded from 64 active electrodes placed on an elastic Acti-Cap according to the 10/20 International System, referenced to the left earlobe. Electrodes mounted on an elastic cap were placed at the following sites: Fp1, Fpz (as ground), Fp2, AF7, AF3, AFz, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT9 (as VEOG), FT7, FC5, FC3, FC1, FCz (as left earlobe reference), FC2, FC4, FC6, FT8, FT10 (as right HEOG), T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP9 (as left HEOG), TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, TP10 (as right earlobe reference), P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, and Iz (as FCz). The EEG was re-referenced offline to the average of the left and right earlobes. Horizontal EOG (HEOG) was recorded as the voltage difference between electrodes placed at the external canthi of the left and right eye, to measure

saccades. To measure blinks, vertical EOG (VEOG) was recorded as the voltage difference between an electrode placed under the left eye and the electrode at Fp1, located above the left eye. The electrode impedance was kept at less than 10 KΩ. EEG and EOG signals were amplified and digitized at a sampling rate of 500 Hz (pass band 0.01–30 Hz and notch filter to remove frequencies between 48 and 52 Hz) and resampled offline at 250 Hz. The EEG was segmented into 1100-ms epochs starting 100 ms before the onset of the memory array. The epochs were baseline-corrected based on the mean activity during the 100-ms pre-stimulus period. Then, ocular artifacts rejection was performed in two different ways for the same data (see section below) and epochs with artifacts non-eye movement-related (signal exceeding ±100 μV within 1 second time window after memory array onset) were excluded from all analyses. Trials associated with incorrect responses were discarded from analysis and we kept participants who showed at least 30 trials in each condition. If a subject had more than 50% of trials removed because of artifacts, was rejected from the analysis. Then, we extracted the lateralized ERPs for each subject, and we created a 'Grand Average' SPCN component considering the channels PO7 and PO8, which are where the SPCN component is maximally expressed (Luria et al., 2016).

5.3.1. Ocular Artifacts rejection and ICA correction

The first approach we used to remove ocular artifacts (horizontal and vertical eye movements) was the traditional epoch-rejection method. In this procedure, epochs exceeding an arbitrary threshold, are labeled, and then excluded from the analysis. The same criteria were established for VEOG and HEOG movements (deflection > 80 μV within 1 second time window after memory array onset). The threshold value used for blinks represents the electrical activity that an average blink produces, and it is largely used in literature (Drisdelle et al., 2017; Meconi et al., 2018). Our goal here was to remove as many blinks as possible from the analysis as artifacts of no interest. The threshold value used for saccades is more lenient and it has been used to leave some saccades in the epochs, for both fixation

and saccade blocks. This allows us to evaluate the efficacy of ICA correction when subjects make saccades by mistake under fixation instructions, and the effects of ICA and saccades on the lateralized ERP when subjects perform a lot of lateral eye movements during the task.

In the second approach, we performed ICA on the segmented data for each subject to identify and subsequently remove ocular components. In this manner, we can save epochs for the analysis by removing only the specific artifact from the data, instead of removing the whole epoch, and so, useful data. We used the infomax (information maximization) ICA algorithm, developed by Bell and Sejnowski (1995), and then we removed manually only the artifacts representing ocular artifacts. In a few cases, components representing obvious muscular artifacts or channel noises were discarded. After ICA correction, we also performed the epoch-rejection procedure on the 'clean' data with the same criteria used in the first approach. We do that to remove any big artifacts left in the data and to evaluate the number of epochs saved performing ICA compared to the epoch-rejection method.

5.4. Procedure and Statistical Analysis

When participants arrived at the lab, we gave them the formal consent to read and sign. After that, we started to measure the participants' head and scrub the areas where EOG electrodes must be placed. We put an elastic cap on their head, filled the electrodes' holes with electroconductive gel before attaching each electrode in the right position, and then we checked the electrodes' impedance. On average, this procedure took 45 minutes. We explained to the subjects the task they were about to do, instructing them to avoid blinking before they gave the response (each trial lasted around 5 seconds). Then, we left them alone to do the experiment that lasted on average 1 hour and 15 minutes.

The Data processing was performed using custom MATLAB code that called functions in the EEGLAB, MATLAB toolbox, and the ERPLAB plugin for EEGLAB. The statistical analyses for EEG data were conducted using Rstudio on the mean values of SPCN components for each condition

and participant, taken within the time interval of interest. We also used Rstudio to analyse the accuracy and Cowan's K. The Cowan' K is an estimate of working memory capacity extracted using change detection paradigms (Figure 10), representing the number of available slots at each set size (Cowan, 2001; Pashler, 1988; Rouder et al., 2011). It comes from a popular conceptualization that considers the visual working memory being composed of a limited number of slots, one for each item to remember (Cowan, 2001). The formula is:

$$K = S * (H - F)$$

where K is the memory capacity, S is the size of the array, H is the observed hit rate, and F is the false alarm rate (Vogel & Machizawa, 2004). We used Cowan's formula instead of Pashler's formula because it seems more appropriate when we have a single-probe change detection task (Rouder et al., 2011).

6. RESULTS

6.1. Accuracy data

The accuracy results obtained from the subjects kept for the final analysis can be seen in Figure 11. Notably, the accuracy for the fixation block was 77.74% (± 11.71%), and for the free gaze/saccade block was 82.11% (± 10.89%). For the different set sizes (2, 3, 5), the accuracy was respectively 88.44% (± 8.44%), 80.91% (± 9.50%), and 70.43% (± 8.55%). When the presentation of the memory array lasted 100 ms the subjects obtained a 77.00% (± 11.82%) accuracy, whereas with the presentation time at 500 ms, the accuracy was 82.85% (± 10.43%). These results suggest that subjects were more accurate when they could saccade toward the cued side of the monitor and when the presentation time was longer (500 ms). Additionally, as the set size of the memory array increased, accuracy decreased.

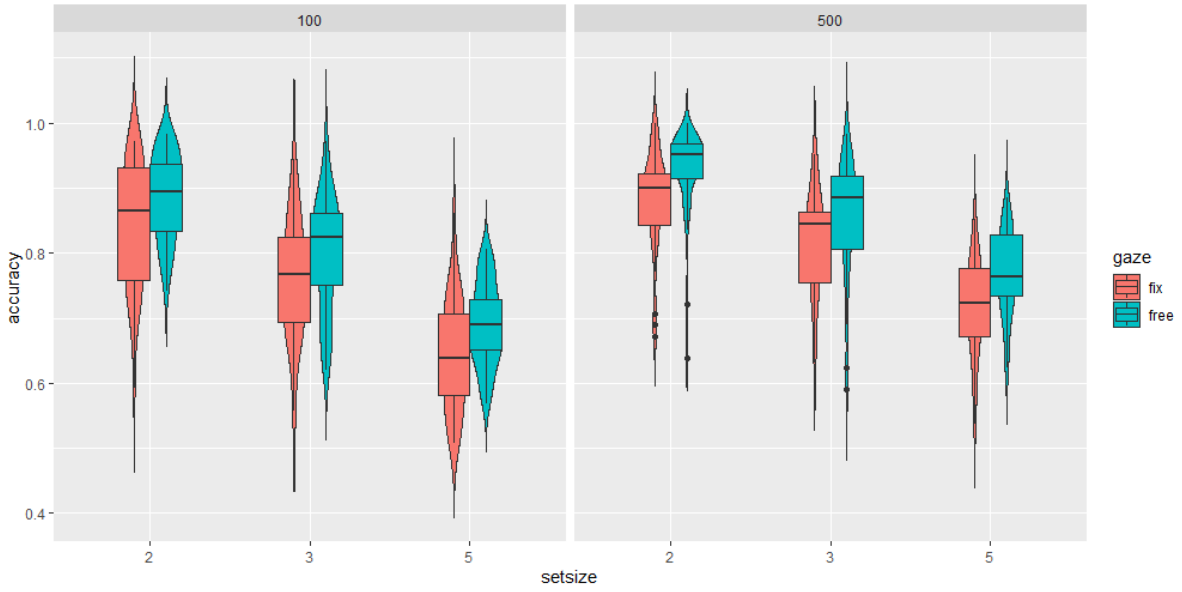


Figure 11: Violin boxplot representing the mean subjects' accuracy at different memory set size, divided for experimental blocks and presentation time.

To assess the effect of these variables, we conducted a 2 x 2 x 2 x 3 mixed ANOVA, examining the different blocks (within-subject, two levels: fixation or free gaze/saccade), the presentation time (between-subjects, two levels: 100 ms or 500 ms), the side of appearance of the memory array (within-

subject, two levels: left or right) and the set size of the memory array (within-subject, three levels: 2, 3, or 5). A significant difference was observed between: fixation and saccade blocks, $F(1,34) = 23.71$, $p < 0.001$; presentation times of the memory array, $F(1,34) = 8.98$, $p = 0.005$; and different set sizes, $F(2,68) = 339.76$, $p < 0.001$. No significant difference in accuracy was observed between the left and right presentation side, $F(1,34) = 0.0005$, $p = 0.98$. No interactions between experimental factors were significant (see Table 1). The between-subjects factor 'presentation' had different variances between the two groups, but when it was corrected with the Welch test, it remained significant ($p < .001$).

	Effect	DFn	DFd	F	p	p<.05	ges
2	presentation	1	34	8.9821784	0.0050641	*	0.1197211
3	setsize	2	68	339.7585848	0.0000000	*	0.4646149
5	gaze	1	34	23.7067053	0.0000254	*	0.0708659
7	side	1	34	0.0005460	0.9814940		0.0000006
4	presentation:setsize	2	68	2.9191181	0.0607769		0.0074009
6	presentation:gaze	1	34	0.0028131	0.9580115		0.0000091
8	presentation:side	1	34	4.0234919	0.0528802		0.0042043
9	setsize:gaze	2	68	0.2846589	0.7531628		0.0005181
11	setsize:side	2	68	0.6539564	0.5232234		0.0014949
13	gaze:side	1	34	2.6747485	0.1111772		0.0036872
10	presentation:setsize:gaze	2	68	0.1531242	0.8583182		0.0002787
12	presentation:setsize:side	2	68	1.1579124	0.3202569		0.0026438
14	presentation:gaze:side	1	34	1.0808567	0.3058431		0.0014933
15	setsize:gaze:side	2	68	0.5812992	0.5619255		0.0011355
16	presentation:setsize:gaze:side	2	68	0.6981418	0.5010393		0.0013634

Table 1: Four ways ANOVA results for accuracy data.

Thus, performance is affected by experimental factors, and accuracy is higher when: the set size of the memory array is smaller; the subjects can saccade toward the cued side; and the presentation time of the memory array is longer (500 ms).

Post-hoc comparisons were conducted via three dependent t-tests using the false discovery rate (FDR; Benjamini & Hochberg, 1995) correction for multiple comparisons. The accuracy with set size 2 was significantly higher than accuracy with set size 3, $t(143) = 12.67$, $p < .001$; accuracy with set size 3 was higher than accuracy with set size 5, $t(143) = 15.98$, $p < .001$; and accuracy with set size 2 resulted higher than accuracy with set size 5, $t(143) = 27.31$, $p < .001$.

6.2. Cowan's K

The results for 'Cowan's K' from the subjects included in the final analysis are presented in Figure 12. The K value for the fixation block was 1.67 (± 0.65) and for the free gaze/saccade block was 1.96 (± 0.62). The Ks for the different set sizes (2, 3, 5) corresponded respectively to 1.54 (± 0.34), 1.85 (± 0.56), and 2.05 (± 0.85). When the presentation of the memory array lasted 100 ms the K was 1.60 (± 0.60), whereas with the presentation time at 500 ms, the K was 2.02 (± 0.64). The image suggests that there could be differences in the estimated K between fixation and saccade blocks, different set sizes of the memory array, and presentation times.

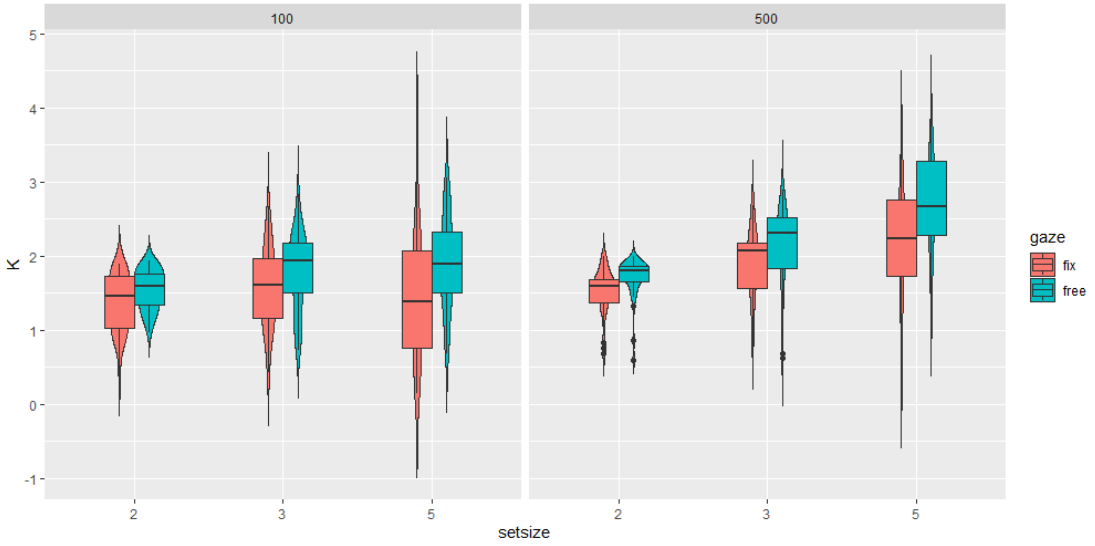


Figure 12: Violin boxplot representing the mean K values at different set sizes, divided for experimental blocks and presentation times.

To test these observations, we conducted a 2 x 2 x 2 x 3 mixed ANOVA, like for the accuracy analysis. A significant difference was present between: fixation and saccade blocks, $F(1,34) = 24.04$, $p < 0.001$; presentation times, $F(1,34) = 11.44$, $p = 0.002$; and the different set sizes of the memory array, $F(2,68) = 46.55$, $p < 0.001$. Besides, two interactions were observed: one between presentation time and the set size of the memory array, $F(2,68) = 15.33$, $p < 0.001$; and another between the different experimental blocks (fixation and free gaze/saccade block) and the set sizes, $F(2,68) = 6.11$, $p = 0.004$. No other main effects and interactions were significant (see Table 2). The between-subjects

factor ‘presentation’ had different variances between the two groups, but when was corrected with the Welch test, it remained significant ($p < .001$).

	Effect	DFn	DFd	F	p	p<.05	ges
2	presentation	1	34	11.4441605	0.0018190	*	0.1325307
3	setsize	2	68	46.5461393	0.0000000	*	0.1308591
5	gaze	1	34	24.0377795	0.0000229	*	0.0691346
7	side	1	34	0.1724202	0.6805786		0.0002317
4	presentation:setsize	2	68	15.3259330	0.0000032	*	0.0472328
6	presentation:gaze	1	34	0.0027764	0.9582859		0.0000086
8	presentation:side	1	34	2.4305914	0.1282488		0.0032560
9	setsize:gaze	2	68	6.1051071	0.0036433	*	0.0139317
11	setsize:side	2	68	0.9095799	0.4075360		0.0022806
13	gaze:side	1	34	3.5123854	0.0695260		0.0048644
10	presentation:setsize:gaze	2	68	0.1244931	0.8831452		0.0002880
12	presentation:setsize:side	2	68	1.8092569	0.1715705		0.0045261
14	presentation:gaze:side	1	34	2.0630938	0.1600417		0.0028630
15	setsize:gaze:side	2	68	1.3621662	0.2630037		0.0029534
16	presentation:setsize:gaze:side	2	68	1.9051269	0.1566619		0.0041258

Table 2: Four ways ANOVA results for Cowan’s K.

Thus, Cowan’s K is influenced by experimental factors, and it was higher when: the presentation time of the memory array was longer (500 ms); the set size of the memory array was bigger; and the subjects could saccade toward the cued side. However, the differences in the K’s values for different set sizes changed at different presentation times of the memory array and between the two experimental blocks (fixation and saccade block).

In a post hoc analysis using the false discovery rate method, we found that: the K with set size 2 was significantly smaller than the value of K with set size 3, $t(143) = -9.15$, $p < .001$; the value of K with set size 3 was smaller relative to the K with set size 5, $t(143) = -3.46$, $p = 0.004$; and the value of K with set size 2 was smaller than the value of K with set size 5, $t(143) = -8.54$, $p < 0.001$. Moreover, with presentation time of 100 ms: the value of K with set size 2 was significantly smaller than the value of K with set size 3, $t(71) = -4.27$, $p < .001$; the value of K with set size 3 was not different from the value of K with set size 5, $t(71) = -0.01$, $p = 1$, as well as the value of K between set size 2 and 5, $t(71) = -2.71$, $p = 0.06$. However, with a presentation time of 500 ms: the value of K with set size 2 was smaller than the value of K with set size 3, $t(71) = -9.43$, $p < .001$; the value of K with set

size 3 was significantly smaller than the value of K with set size 5, $t(71) = -5.62$, $p < .001$; and the value of K with set size 2 was smaller than the value of K with set size 5, $t(71) = -10.86$, $p < .001$.

Looking at the previous results, we can observe that the differences between K values at different set sizes change at different presentation times, and this explains the significant interaction we found between presentation times and set sizes. At 100 ms the K values at set sizes 3 and 5, as well as the Ks at set sizes 2 and 5, are almost the same; whereas with 500 ms there is an increase of the Ks as the set size increases.

For what concerns the interaction between experimental blocks and set sizes of the memory array, in the fixation block of the experiment: the value of K with set size 2 was significantly smaller than the value of K with set size 3, $t(71) = -5.73$, $p < .001$; the value of K with set size 3 was not different from the value of K with set size 5, $t(71) = -0.88$, $p = 1$; and the value of K with set size 2 was smaller than the value of K with set size 5, $t(71) = -4.10$, $p < .001$. However, in the free gaze/saccade block: the value of K with set size 2 was smaller than the value of K with set size 3, $t(71) = -7.25$, $p < .001$; the value of K with set size 3 was significantly smaller than the value of K with set size 5, $t(71) = -4.32$, $p < .001$; and the value of K with set size 2 was smaller than the value of K with set size 5, $t(71) = -8.53$, $p < .001$. As we can see, the differences between K values at different set sizes change whether subjects make or not a saccade, and this explains the interaction between set sizes and the experimental blocks. When subjects fixate on the center of the monitor, the K values between set sizes 3 and 5 are the same, while in the saccade condition, there is an increase of the K as the set size increases (see Figure 13). All the post-hoc analyses were conducted via dependent t-tests using the false discovery rate to correct for all the comparisons performed.

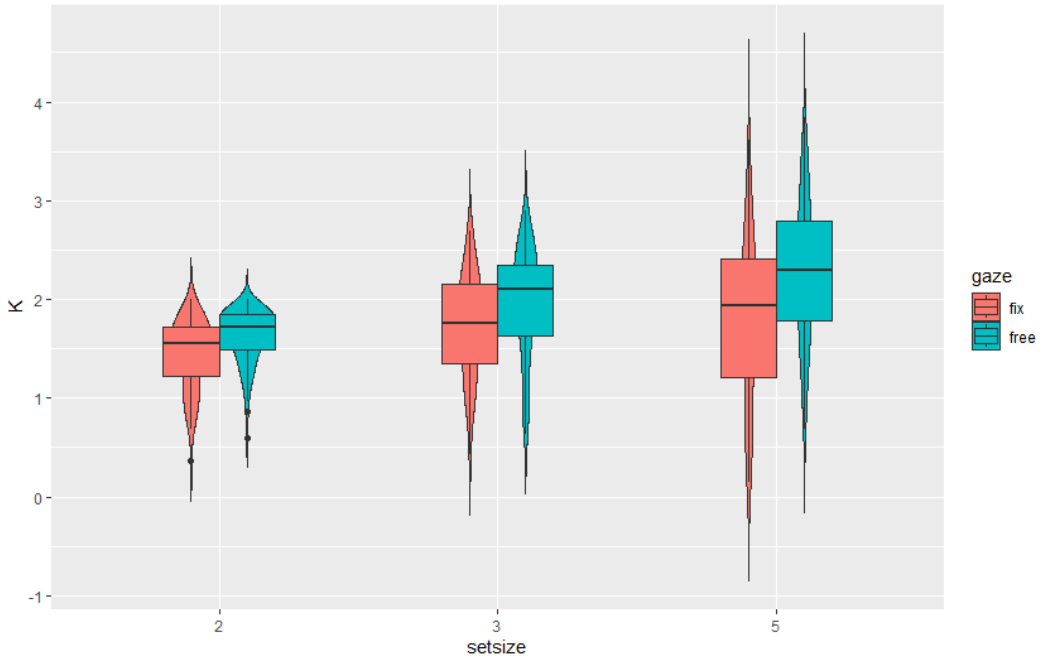


Figure 13: Violin boxplot representing mean K values at each set size, divided for fixation and saccade blocks.

6.3. EEG DATA

6.3.1. Trials rejection

As already mentioned, an important concern with the epoch-rejection approach is the possible small number of non-artefactual trials to extract the ERP averages. In this experiment, we analyzed the same EEG data using two different methods to evaluate the differences in the number of trials saved and the final ERP averages extracted. For both methods of analysis, we used the same rejection criteria for HEOG and VEOG (change of at least 80 µV within a 1 second time window from the stimulus onset). Then, we also checked for artifacts higher than 100 µV in all the other channels.

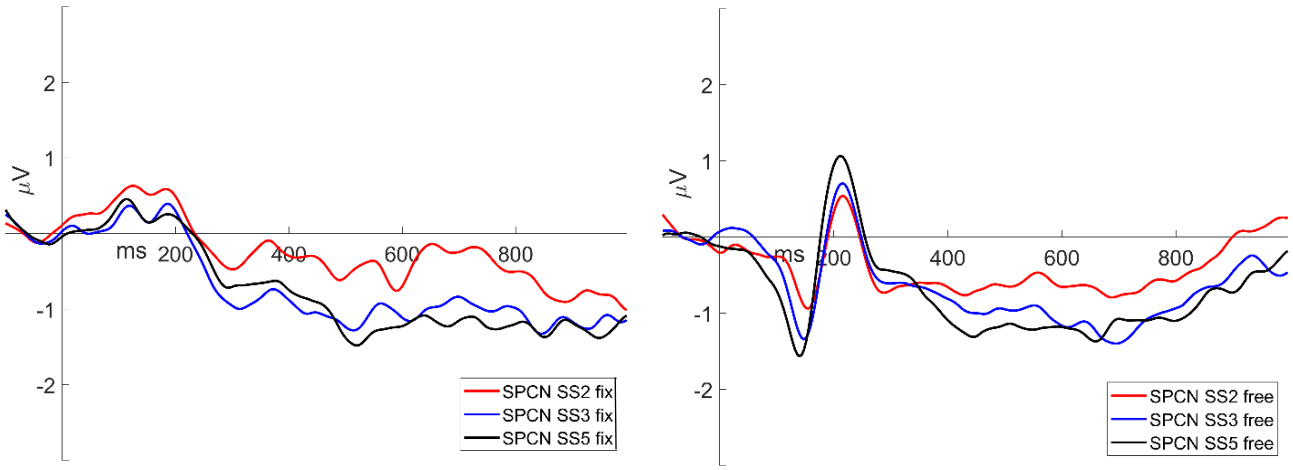
Thus, for the data with the presentation time at 100 ms and epoch-rejection approach, 6.07% of the epochs (SD = 6.22%) were labeled as containing ocular artifacts (blinks or saccades), and the rejections due to saccades were 1.71%, (SD = 4.53%) of the total artifacts. After ICA correction, instead, 0.75% of the epochs (SD = 1.06%) were labeled as containing ocular artifacts (blinks or saccades), and the rejections due to saccades were 0.37%, (SD = 0.89%) of the total artifacts. For the

data with a presentation time of the memory array of 500 ms and epoch-rejection approach, 5.64% of the epochs (SD = 6.07%) were labeled as containing ocular artifacts (blinks or saccades), and the rejections due to saccades were 1.01%, (SD = 2.10%) of the total artifacts. After ICA correction, 0.24% of the epochs (SD = 0.31%) were labeled as containing ocular artifacts (blinks or saccades), and the rejections due to saccades were 0.04%, (SD = 0.10%) of the total artifacts. Overall, when the ICA method was used in addition to the epoch-rejection approach, 85% of epochs labeled as ocular artifacts, and in particular 81% of epochs with saccades, were kept for the analysis.

6.1.2 SPCN amplitude

Exploratory analysis

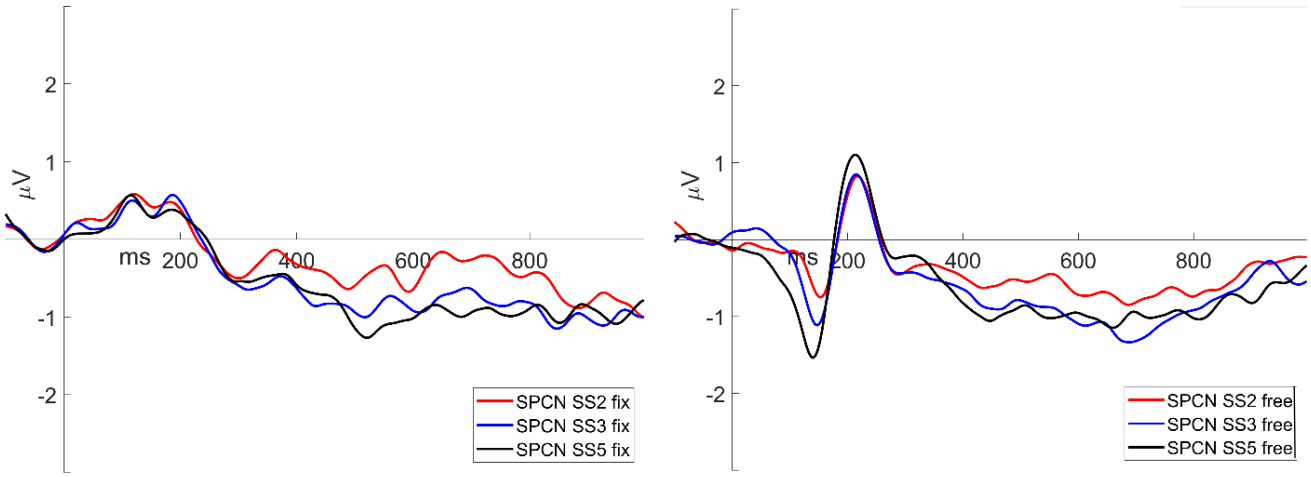
After conducting the preprocessing to extract the SPCN components, we obtained several SPCN separated by set size of the memory array, experimental block (fixation and free gaze/saccade blocks), and presentation time. Figures 14 and 15 illustrate the lateralized ERPs at a presentation time of 100 ms, following the epoch-rejection method. Figure 14 shows SPCN components for each set size in the fixation blocks, while Figure 15 represents the components in the free gaze/saccade blocks.



Grand-averaged SPCN waveforms (at PO7/PO8) at 100 ms of presentation time for the fixation condition after the epoch-rejection procedure on the left (Figure 14), and for the saccade condition after the epoch-rejection procedure on the right (Figure 15).

Looking at the images, we can appreciate the SPCN components from 300 ms post-stimulus onset until the end of the temporal window (1 second after stimulus onset). It seems there is an effect of the set size of the memory array, with increased negativity as the set size increases (reaching a plateau around a set size of 3, given that with 3 and 5 elements the waves are very similar). It seems that in the saccade blocks, there is a decrease in negativity at the end of the temporal window.

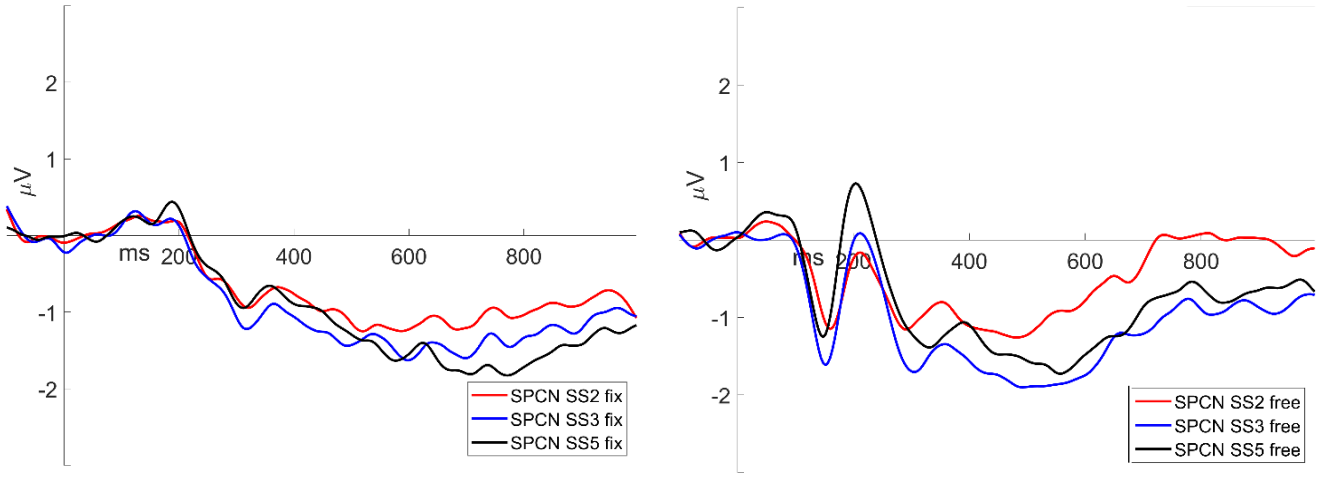
Instead, Figures 16 and 17 illustrate the SPCN components for each set size at a presentation time of 100 ms following ICA correction, in the fixation and the free gaze/saccade blocks, respectively.



Grand-averaged SPCN waveforms (at PO7/PO8) at 500 ms presentation time for the fixation condition after ICA correction on the left (Figure 16), and for the free gaze/saccade condition after ICA correction on the right (Figure 17).

These images are very similar to those obtained with the epoch-rejection method, even though it seems that the differences between the components at different set sizes are reduced, as well as the decrease in negativity at the end of the temporal window for SPCN in the saccade blocks.

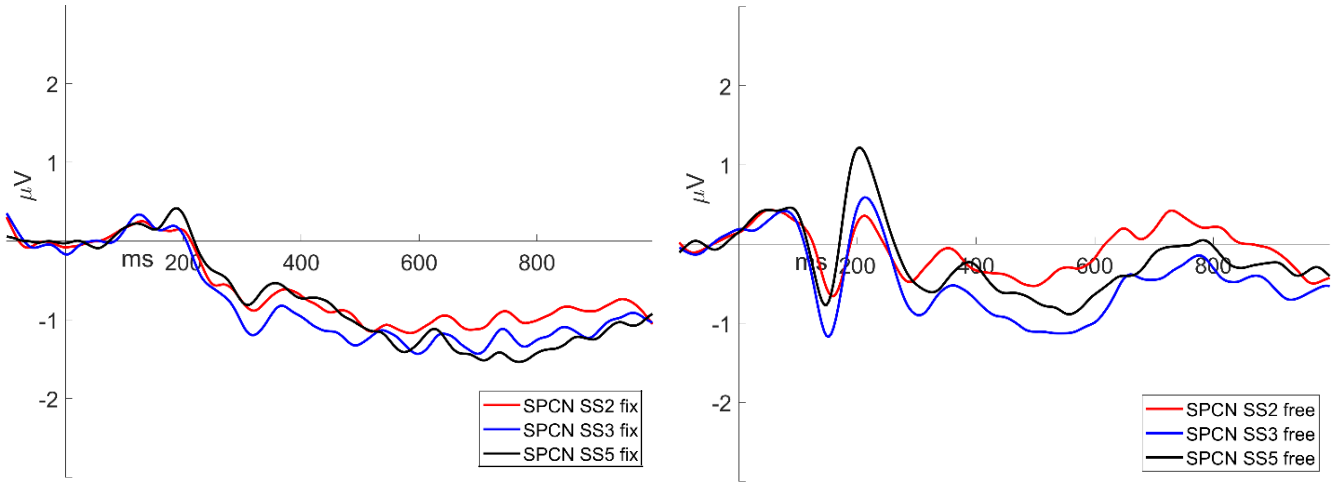
Figures 18 and 19 represent the SPCN components for each set size at a presentation time of 500 ms following the epoch-rejection method, in the fixation and the free gaze/saccade blocks, respectively.



Grand-averaged SPCN waveforms (at PO7/PO8) at 500 ms of presentation time for the fixation condition after the epoch-rejection procedure on the left (Figure 18), and for the saccade condition after the epoch-rejection procedure on the right (Figure 19).

We can appreciate how the increase in negativity for larger set sizes seems to be maintained. Probably the amplitude of the components is more negative compared to those for the presentation time of 100 ms, and the decrease in negativity at the end of the temporal window for the saccade condition is even more evident.

Finally, we have the last two images (Figures 20 and 21) illustrating the SPCN components for each set size at a presentation time of 500 ms following ICA correction, in the fixation and free gaze/saccade blocks, respectively.



Grand-averaged SPCN waveforms (at PO7/PO8) at 500 ms presentation time for the fixation condition after ICA correction on the left (Figure 20), and for the free gaze/saccade condition after ICA correction on the right (Figure 21).

The differences between the components at different set sizes seem reduced as well as the amplitude (especially in the free gaze condition). The decrease in negativity in the blocks where the subjects foveated the cued side of the monitor is still present.

Statistical analysis

To test our hypotheses, we conducted the statistical analyses in Rstudio using the package "ez" (Lawrence & Lawrence, 2016). SPCN amplitude was measured making the mean of the time points within a time window of 500 ms. This time window started at 500 ms post-stimulus onset and ended at the end of the epochs (1 second post-stimulus onset). We chose this interval because, when considering a classical time window beginning at 300 ms post-stimulus onset, in the condition with a presentation time of 500 ms we would also take into account 200 ms during which the participants still had the memory array present on the screen. For this reason, we reduced the time interval of interest considering only the SPCN amplitude during the retention period after the memory array presentation.

Then, we performed a 2 x 2 x 2 x 3 mixed ANOVA, examining the method of analysis (within-subject, 2 levels: epoch-rejection and ICA correction), the experimental blocks (within-subject, 2 levels: fixation block and free gaze/saccade block), the presentation time of the memory array (between-subjects, 2 levels: 100 ms and 500 ms) and the set size of the memory array (within-subject, 3 levels: 2, 3, and 5). The results obtained can be observed in Table 3.

	Effect	DFn	DFd	F	p	p<.05	ges
2	presentation	1	34	0.0381024	0.8463994		0.0005871
3	ICA	1	34	13.2234680	0.0009055	*	0.0028892
5	setsize	2	68	8.9597619	0.0003517	*	0.0155704
7	gaze	1	34	1.4169554	0.2421504		0.0126583
4	presentation:ICA	1	34	5.5521523	0.0243615	*	0.0012151
6	presentation:setsize	2	68	0.0951496	0.9093578		0.0001679
8	presentation:gaze	1	34	0.6574031	0.4231177		0.0059130
9	ICA:setsize	2	68	6.8098822	0.0020149	*	0.0010203
11	ICA:gaze	1	34	0.8847211	0.3535470		0.0001104
13	setsize:gaze	2	68	0.3010396	0.7410296		0.0007774
10	presentation:ICA:setsize	2	68	0.0387260	0.9620354		0.0000058
12	presentation:ICA:gaze	1	34	11.3702130	0.0018739	*	0.0014167
14	presentation:setsize:gaze	2	68	0.3623332	0.6973861		0.0009355
15	ICA:setsize:gaze	2	68	0.5125669	0.6012517		0.0000533
16	presentation:ICA:setsize:gaze	2	68	0.5654075	0.5707764		0.0000588

Table 3: Four ways ANOVA for EEG data.

A significant difference was observed between methods of analysis, $F(1,34) = 13.22$, $p < 0.001$, and set sizes of the memory array, $F(2,68) = 8.96$, $p < 0.001$. Two interactions were observed between: the presentation times of the memory array and the methods of analysis, $F(1,34) = 5.55$, $p = 0.02$; and methods of analysis and set sizes of the memory array, $F(2,68) = 6.81$, $p = 0.004$ (corrected for sphericity after Mauncly's test). Besides, a three-way interaction was observed between presentation times of the memory array, methods of analysis, and experimental blocks, $F(1,34) = 11.37$, $p = 0.002$. Table 4 and Figure 22 can help us to interpret better the significant main effects.

Condition	Levels	Analysis	SPCN_amplitude
Set size	2	ICA	-0.55 (±1.82) µV
		NoICA	-0.59 (±1.74) µV
	3	ICA	-0.92 (±1.72) µV
		NoICA	-1.15 (±1.64) µV
	5	ICA	-0.91 (±1.96) µV
		NoICA	-1.21 (±2.03) µV
Gaze	Fixation	ICA	-1.02 (±1.62) µV
		NoICA	-1.17 (±1.62) µV
	Free gaze	ICA	-0.57 (±2.01) µV
		NoICA	-0.80 (±2.00) µV
Presentation	100 ms	ICA	-0.81 (±2.06) µV
		NoICA	-0.88 (±2.03) µV
	500 ms	ICA	-0.77 (±1.58) µV
		NoICA	-1.09 (±1.60) µV

Table 4: The SPCN mean amplitudes for the three factors (set size, experimental blocks, and presentation time) at different methods of analysis (epoch-rejection method and ICA correction). Between brackets are reported the standard deviations.

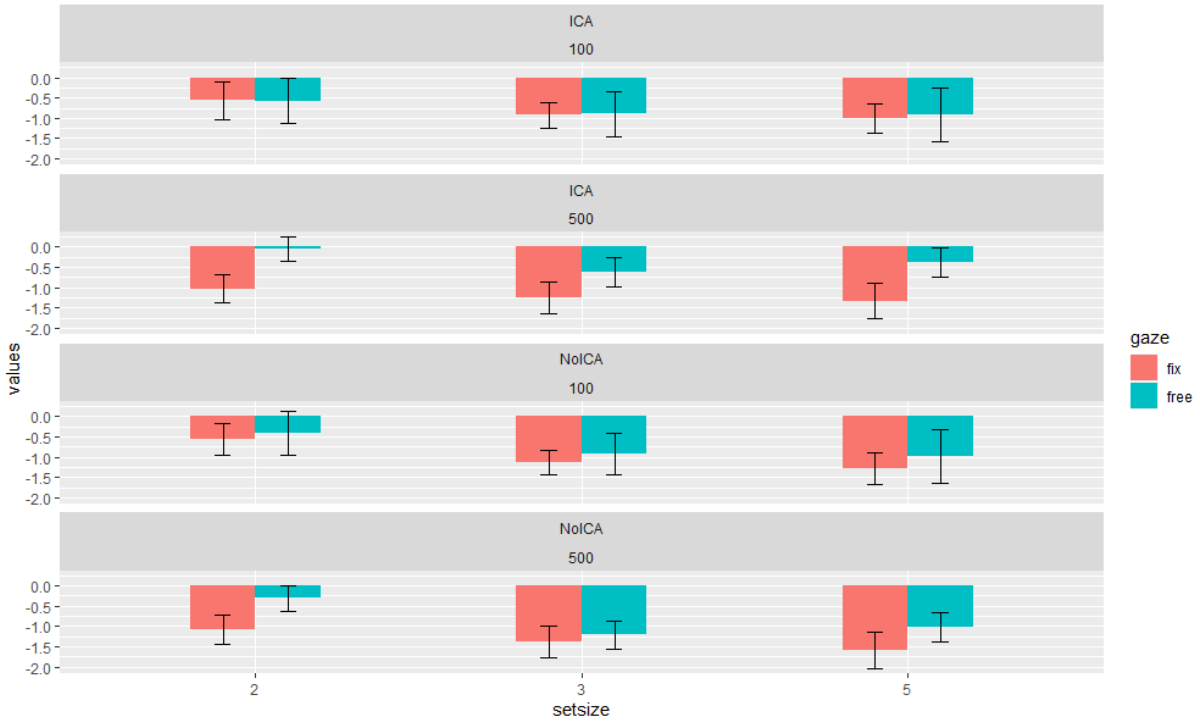


Figure 22: SPCN mean amplitudes at each specific condition.

As we can appreciate also from Table 4, the amplitudes of the SPCN, when the ICA correction was performed, were lower almost everywhere compared to the amplitudes of the components under the epoch-rejection method. For what concerns the results of the set size corrected for multiple comparisons with the false discovery rate method of Benjamini and Hochberg, there was: a significant difference between the memory array with the set size of 2 and 3, $t(143) = 4.81$, $p < .001$; and a significant difference between the memory array with the set size of 2 and 5, $t(143) = 4.21$, $p < .001$. No difference was observed between the memory array with set sizes 3 and 5. Overall, we can say that when the memory array has set size 2, the SPCN amplitude is more positive than with set sizes 3 and 5. It is interesting to notice that overall, SPCN amplitude is not affected by both different presentation times of the memory array and experimental blocks (fixation and saccade blocks).

To understand better the significant interactions, some post-hoc analyses and graphs might be helpful. Figure 23 shows the interaction between the presentation times and the methods of analysis.

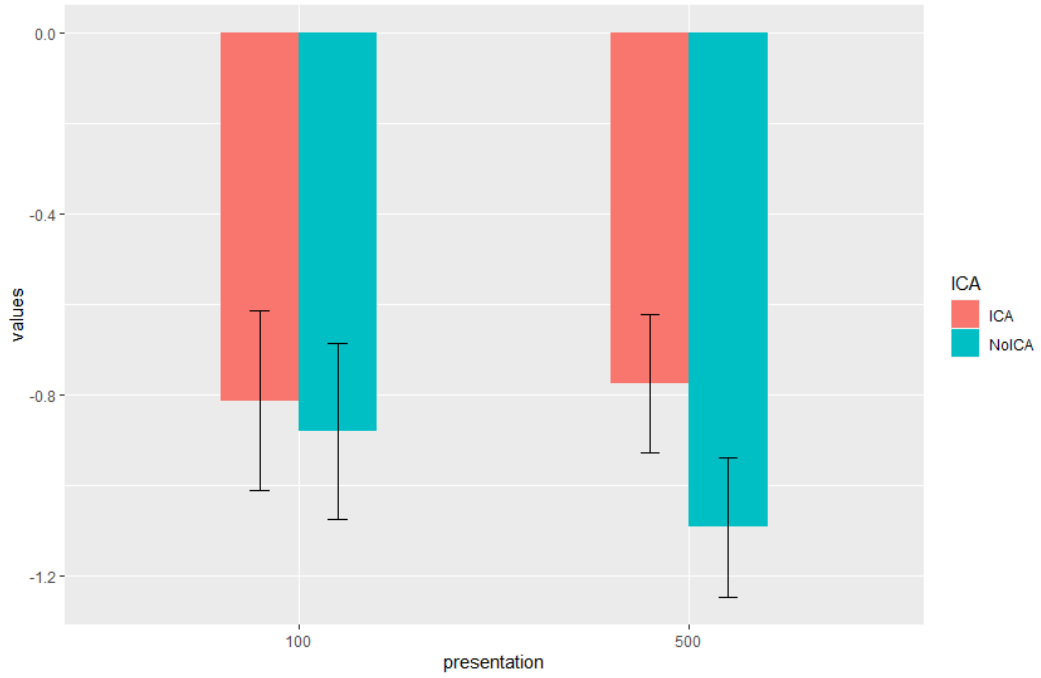


Figure 23: SPCN amplitude values for different methods of analysis and presentation times.

This interaction indicates that there is a difference in the SPCN amplitudes between the methods of analysis, across the levels of presentation time of the memory array. After conducting a post-hoc analysis, and correcting for multiple comparisons with the false discovery rate method, it appeared that SPCN amplitude is almost the same with a presentation time of 100 ms between the two methods of analysis, but there was a significant difference when the presentation time is of 500 ms, $t(107) = -6.07$, $p < 0.001$: the data analyzed with epoch-rejection method only, are more negative than those analyzed also with ICA correction when the presentation time is longer. The differences between the same method of analysis across the levels of the presentation time were not significant.

The other interaction is between methods of analysis and set sizes of the memory array. As we can see also from Figure 24, when ICA was performed, only the SPCN amplitude at set size 2 was more positive than that at set size 3, $t(71) = 2.79$, $p = 0.05$; all other comparisons were not significantly different. However, when the epoch-rejection method was performed, the amplitude at set size 2 was significative more positive than with set size 3, $t(71) = 3.98$, $p = 0.003$, and the amplitude with set size 2 was more positive than with set size 5, $t(71) = 3.82$, $p = 0.003$. To be precise, it was not the

amplitude at set size 2 to be more positive, but those of set size 3 and 5 to be more negative with epoch-rejection approach than with ICA correction ($t(71) = -3.87$, $p = 0.003$; $t(71) = -4.58$, $p < .001$, respectively). Thus, with both methods of analysis, there is no difference between set sizes 3 and 5, but with ICA correction set sizes 2 and 5 do not differ.

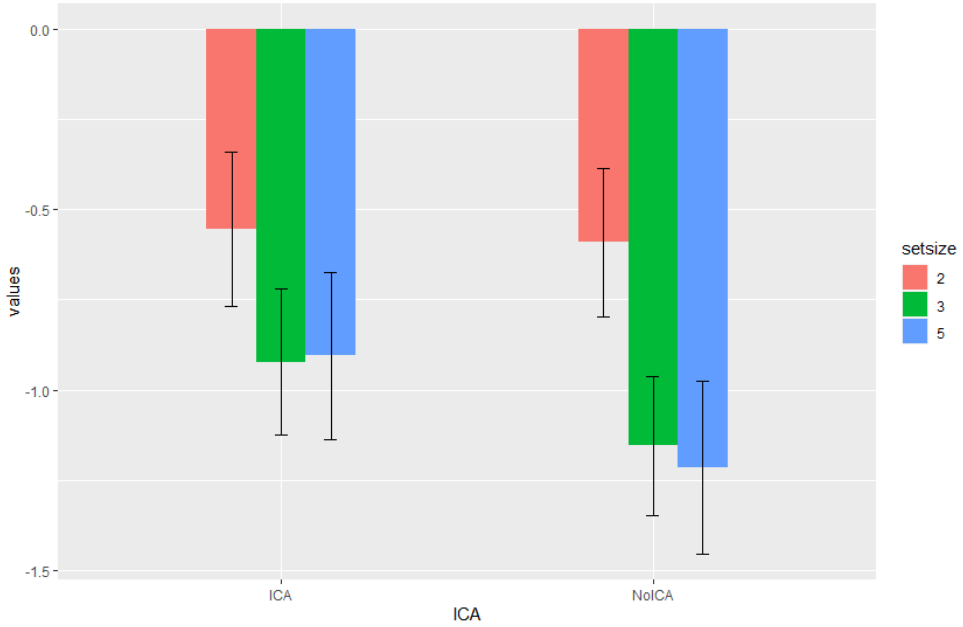


Figure 24: The SPCN mean values at different set size for the two methods of analysis.

For what concerns the three-way interaction between presentation times, methods of analysis, and experimental blocks (fixation and free gaze/saccade), Figure 25 and Table 5 can help us. This interaction means that the interaction between the methods of analysis and the experimental blocks varies across the levels of the presentation time.

Presentation	Gaze	Analysis	SPCN_amplitude
100 ms	Fixation	ICA	-0.82 (±1.60) µV
		NoICA	-0.99 (±1.54) µV
	Free gaze	ICA	-0.80 (±2.46) µV
		NoICA	-0.77 (±2.43) µV
500 ms	Fixation	ICA	-1.20 (±1.63) µV
		NoICA	-1.35 (±1.70) µV
	Free gaze	ICA	-0.35 (±1.43) µV
		NoICA	-0.84 (±1.47) µV

Table 5: SPCN mean amplitudes in each specific condition of the three-way interaction.

In the post-hoc analyses, the comparisons, corrected with the false discovery rate method, in the condition with presentation time 100 ms between experimental blocks and methods of analysis, resulted not significant: basically, there was not a reliable difference in the SPCN values between them. Anyway, in the condition with a presentation time of 500 ms, when participants made a saccade toward the cued side, SPCN amplitude was more positive when ICA correction was performed ($t(53) = 5.96$, $p < 0.001$, respectively).

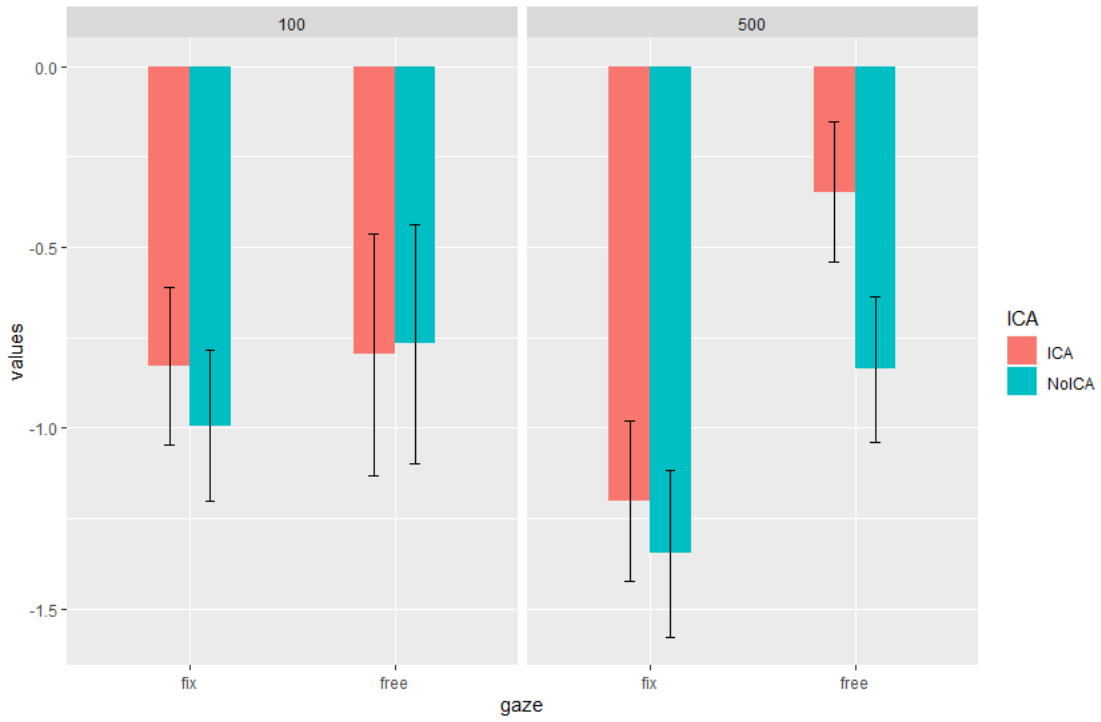


Figure 25: SPCN amplitude values for each level of the factors of the three-way interaction.

7. DISCUSSION

With this experiment, we wanted to evaluate three main points: whether ICA correction of ocular artifacts (blinks and, in particular, saccades) could be used as a reliable method to avoid high numbers of epochs rejected in visual working memory experiments studying lateralized ERPs (SPCN component); whether letting subjects to make a saccade toward lateralized items during the encoding phase of the task, could influence the SPCN waveform; and whether experimental paradigm with a longer presentation time of the memory array and less constraints on the eye movements, could produce similar or better results compared to the traditional design. Using a mixed-subject design, we assessed the presence of the memory load effect on the SPCN components under two different methods of data analysis (epoch-rejection method and ICA correction) when participants were instructed to maintain fixation at the center of the monitor, and also when they were instructed to saccade toward target items at the side of the monitor. Besides, two groups of subjects conducted two versions of the same experiment (one with a presentation time of the memory array of 100 ms and a retention time of 900 ms, and another with a presentation time of 500 ms and a retention time of 500 ms) to evaluate whether, with more time to efficiently saccade toward the items, there could be any differences in the SPCN components compared to the experiment with less time.

The study produced several interesting results. First of all, we observed that with ICA, more than 80% of the epochs removed using the epoch-rejection approach were saved. There was also a consistent reduction in the variability of discarded epochs among subjects and conditions. We succeeded in replicating the effect of memory load on the SPCN amplitude reported in literature (Luria et al., 2016; Vogel & Machizawa, 2004), with increased negativity observed as the number of representations being held in memory increased. We obtained this finding for both experiments with different presentation times, noticing that SPCN amplitude reached an asymptotic limit at set size 3 in the 'Grand Average' ERP. Then, we observed that the method of analysis used influenced the SPCN amplitude, with increased negativity for the ERPs following the epoch-rejection approach. This result

was expected since the trials' rejection was done using a lenient rejection criterion that led to the creation of saccade-conserved data for the condition in which subjects could move their eyes toward the target items. Thus, the reduced amplitude negativity with the ICA correction method may be due to the removal of the saccades' activity. The same reason might be the cause of the smaller amplitude differences between set sizes after performing the ICA correction (set size 5 did not differ from set size 2). In contrast, with the epoch-rejection method, the differences were greater and the amplitudes at set sizes 3 and 5 were more negative. Perhaps, the SPCN waves were not affected at set size 2 because, with so few items to memorize, subjects did not perform large saccades (this hypothesis must be tested). However, we saw in the three-way interaction, that these differences caused by the method of analysis used, were mainly driven by the experiment with a presentation time of 500 ms. In this condition, there was a big reduction in the SPCN amplitude after the ICA correction, while in the experiment with a presentation time of 100 ms, it was negligible. Indeed, we observed that, only when the presentation time was longer, the SPCN amplitude was reduced after ICA, especially when subjects had to make a saccade.

An explanation for these results may be that subjects with the longer presentation time and the possibility to saccade the target items have more time to make eye movements and navigate the space where the items are (given that the elements are still present on the monitor when they complete the saccade). This aspect may have resulted in increased saccade activity in this condition compared to when subjects were required to saccade with a presentation time of 100 ms. This could have led to a significant reduction in amplitude after ICA correction due to the removal of large artifacts, and perhaps an increased difficulty in disentangling saccades from the neural signal. Besides, the fact that also with a presentation time of 500 ms in the fixation blocks, there was a reduction in SPCN amplitude following ICA correction (even if it was not significant after the multiple comparisons correction), it may indicate that subjects were more inclined to break the fixation instruction when the presentation time was longer. In short, even though the presentation time does not result significant in the first statistical analysis, there might be some specific situations in which a longer

presentation time, interacting with other factors (method of analysis and ocular movements) might generate unreliable or problematic SPCN components. However, it is important to pinpoint that in experiments with a presentation time of 100 ms, there were no significant alterations or differences due to different methods of analysis or eye movements, and the SPCN values at different set sizes under any conditions were similar (except for a small reduction in negativity after ICA correction). Moreover, it does not seem that making a saccade with a presentation time of 100 ms leads to backward masking: the SPCN components were still present; they did not differ from those of the fixation blocks; and the accuracy and Cowan's K results were not compromised. This suggests that researchers could be more tolerant if subjects do not follow the fixation instruction in each trial of experiments with brief presentation times, given that the ERPs do not disappear and, using ICA correction in conjunction with the epoch-rejection approach produces almost indistinguishable results.

Accuracy and Cowan's K

We also observed interesting results in the behavioral analysis, and it could be useful to interpret these in light of what we observed in the EEG analysis.

For what concerns accuracy, the subjects showed an effect of the factor 'set size'. This was very clear in both experiments with different presentation times, having higher accuracy when the number of item-to-be-remembered was lower. However, when the presentation time was at 500 ms, the accuracy was higher than when it was at 100 ms, and the subjects also had an improvement when they could saccade toward the target items compared to when they had to maintain fixation. The effects of presentation time and experimental condition (fixation and free gaze/saccade blocks) suggest that the less we constrain subjects to follow non-ecological instructions (e.g., maintaining fixation) and the more time we give them to encode the items, the better their performance. Still, the SPCN amplitude with a presentation time of 100 ms does not change between fixation and saccade instructions. This,

perhaps, indicates that the facilitation in the saccade condition does not influence directly this ERP's working memory correlates, but it rather allows the subjects to not perform a second task (e.g., avoiding saccades) and be more focused on the primary task. If we consider the experiment with a presentation time of 500 ms, it may be that the more time the subjects had to accumulate information about the items, the better the final performance. The accuracy increased even more when participants could move their gaze. However, the SPCN amplitudes in these conditions were particularly influenced by these factors yielding results not completely reliable, especially when different methods of analysis were used. It might be that subjects in this condition are free to saccade and navigate the targets' space more effectively (due to the longer presentation time), improving their performance thanks to the more ecological situation, but worsening the EEG data due to larger ocular artifacts. It should be noted that one possible explanation for the improved performance with a presentation time of 500 ms could be that the retention time before the response was 500 ms, compared to 900 ms in the experiment with a presentation time of 100 ms. Besides, differently from us, Drisdelle and colleagues did not observe any differences in accuracy between fixation and saccade blocks, which could potentially be attributed to the use of a different experimental paradigm.

'K' is an estimate of working memory capacity, and it can be considered to reflect the number of available 'slots' for storing distinct objects (Cowan, 2001). The main effects found for the K were the same as those found for accuracy: there was an increase of K values as the set size of the memory array increased; Ks were higher when the presentation time was longer compared to when it was at 100 ms, and they were higher also when subjects could saccade toward the target items. Thus, also here, with a longer presentation time and without a second task (e.g., avoid saccades), the participants were facilitated, and more 'slots' were available. However, in the fixation blocks and in the condition with a presentation time of 100 ms, there was not a significant difference between K values at set sizes 3 and 5. These results are consistent with our observations of the SPCN amplitudes for this condition: at set size 3, both the ERPs' waves and K values reach a plateau. In contrast, with a presentation time of 500 ms, K values continued to increase as the number of items in the memory

array increased, for both the fixation and saccade blocks. In this case, these results contradict what has been found in ERP data, suggesting that this condition is more problematic to handle in EEG studies.

8. CONCLUSION

The current experiment has shown that ICA correction is an effective method for managing ocular activity, particularly saccades, in electrophysiological studies of visual working memory. This is evident from the overall form of SPCN components at different numbers of items presented and the statistical results obtained. Specifically, we replicated the first finding of Drisdelle and colleagues, demonstrating that using ICA to correct trials contaminated with HEOG artifact (saccades), either under fixation or saccade instructions, does not alter the effect of memory load of the SPCN component, previously shown to be involved in VWM (even though slightly reductions in the SPCN amplitude could appear compared to the epoch rejection method). However, this finding is valid only for experiments with briefly presented stimuli (probably less than 200 ms) for which little useful information could be acquired after the saccade. Hence, ICA correction retains a greater proportion of trials than the standard procedure of rejecting trials with artifacts, promoting robust and reliable results without affecting the SPCN properties.

Limitations and Future Directions

A critical aspect of this study was the low effect size associated with most of the results from the ANOVA. One possible explanation could be that the number of subjects per condition was low. Indeed, after removing subjects with accuracy lower than 50% and excessive trials with artifacts, we had only 18 participants in each group (group of the experiment with a presentation time of 100 ms and 500 ms). This number falls below the standard number of subjects, which is typically around 30,

required for cognitive and neurocognitive psychological experiments. Furthermore, it is not based on any objective measure. A 'Bayesian power analysis' could be useful to determine the sample size necessary to strengthen the final effect size. Another reason for the low effect size could be the distribution of the data. In fact, the distribution of the EEG data in this study deviates slightly from the normal distribution necessary to meet the normality assumption of parametric tests, such as ANOVA. With a violation of normality, a statistical test could misrepresent the results or report a low effect size. In this case, it could be useful to replicate the analysis with a non-parametric test like the 'Permutation test', which does not need a normal distribution of the data to be performed and it manages well the 'Type I error'. Another criticality is the experimental design. In the present study the average duration of the experiment was around 2 hours (45/50 minutes for the EEG montage and 1 hour and 15 minutes to complete the task) and a lot of subjects reported tiredness and difficulties remaining focused after 20 minutes from the beginning of the task. Drisdelle and colleagues, in their study, used a multiple-frame procedure (MFP) in which they showed six consecutive visual search presentations, or frames, in each trial (instead of a single one), instructing subjects to respond at the end of these frames the number of frames containing a target. In this way, they presented a larger number of visual stimuli without increasing the duration of the experiment. Trying to apply this procedure to our paradigm could be very useful to improve the quality of the data acquired. Furthermore, some subjects reported difficulty in distinguishing the colors of certain squares. This may have impacted the accuracy of the subjects and, consequently, the extracted ERPs. This issue must be addressed in the future.

For future experiments could be interesting to introduce a third method of analysis, consisting in an epoch-rejection approach with more strict threshold criteria for ocular artifacts. With this approach will be more difficult to reach an appropriate number of trials to extract the ERPs. However, it would be worthwhile to assess the actual number of epochs rejected if we try to remove all the eye movements and the SPCN waveform that it would produce, with the certainty that all the saccades, blinks, and associated cognitive processes have been deleted. Then, to increase the reproducibility,

an automatic procedure to detect and label the ocular artifacts with the ICA correction should be implemented, because manually removing them introduces a high degree of variability due to the differing abilities of the experimenters. We should also investigate whether the observed increase in positivity at the end of the time windows in the SPCN of the saccade condition (Figures 15, 17, 19, 21) could have masked any effects in the means of the SPCN amplitude values used for the statistical analyses. Besides, replicating the SCN component (saccade contralateral negativity) found by Drisdelle and colleagues (Drisdelle et al., 2017), could provide valuable insights into whether it is the cause of the observed increase in positivity.

9. BIBLIOGRAPHY

Becke, A., Müller, N., Vellage, A., Schoenfeld, M. A., & Hopf, J.-M. (2015). Neural sources of visual working memory maintenance in human parietal and ventral extrastriate visual cortex. *NeuroImage*, *110*, 78–86. https://doi.org/10.1016/j.neuroimage.2015.01.059

Bell, A. J., & Sejnowski, T. J. (1995). An Information-Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, *7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berg, P., & Scherg, M. (1991). Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG. *Clinical Physics and Physiological Measurement*, *12*(A), 49. https://doi.org/10.1088/0143-0815/12/A/010

Blinowska, K., & Durka, P. (2006). Electroencephalography (EEG). In *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780471740360.ebs0418

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1999). Affect and the startle reflex. In *Startle modification: Implications for neuroscience, cognitive science, and clinical science* (pp. 157–183). Cambridge University Press. https://doi.org/10.1017/CBO9780511665523.010

Bradley, M. M., & Keil, A. (2012). Event-Related Potentials (ERPs). In V. S. Ramachandran (A c. Di), *Encyclopedia of Human Behavior (Second Edition)* (pp. 79–85). Academic Press. https://doi.org/10.1016/B978-0-12-375000-6.00154-3

Bromm, B., & Treede, R. D. (1984). Nerve fibre discharges, cerebral potentials and sensations induced by CO2 laser stimulation. *Human Neurobiology*, *3*(1), 33–40.

Bugli, C., & Lambert, P. (2007). Comparison between Principal Component Analysis and Independent Component Analysis in Electroencephalograms Modelling. *Biometrical Journal*, *49*(2), 312–327. https://doi.org/10.1002/bimj.200510285

Carpenter, R. H. S. (1988). *Movements of the eyes, 2nd rev. & enlarged ed* (p. 593). Pion Limited.

Chang, W.-D., Cha, H.-S., Kim, K., & Im, C.-H. (2016). Detection of eye blink artifacts from single prefrontal channel electroencephalogram. *Computer Methods and Programs in Biomedicine*, *124*, 19–30. https://doi.org/10.1016/j.cmpb.2015.10.011

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cruccu, G., Aminoff, M. J., Curio, G., Guerit, J. M., Kakigi, R., Mauguiere, F., Rossini, P. M., Treede, R.-D., & Garcia-Larrea, L. (2008). Recommendations for the clinical use of somatosensory-evoked potentials. *Clinical Neurophysiology*, *119*(8), 1705–1719. https://doi.org/10.1016/j.clinph.2008.03.016

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

Dimigen, O., Valsecchi, M., Sommer, W., & Kliegl, R. (2009). Human Microsaccade-Related Visual Brain Responses. *Journal of Neuroscience*, *29*(39), 12321–12331. https://doi.org/10.1523/JNEUROSCI.0911-09.2009

Drisdelle, B. L., Aubin, S., & Jolicoeur, P. (2017). Dealing with ocular artifacts on lateralized ERPs in studies of visual-spatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*, *54*(1), 83–99. https://doi.org/10.1111/psyp.12675

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234. https://doi.org/10.1016/0013-4694(96)95711-9

Fawcett, J., Risko, E., & Kingstone, A. (2015). *The Handbook of Attention*. MIT Press.

Galvani, L. (1791). D viribus electricitatis in motu musculari: Commentarius. In *Bologna: Tip. Istituto delle Scienze*. https://ui.adsabs.harvard.edu/abs/1791dvei.book.....G

Gilchrist, I. (2011). Saccades. In S. P. Liversedge, I. Gilchrist, & S. Everling (A c. Di), *The Oxford Handbook of Eye Movements* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0005

Hu, L., & Zhang, Z. (A c. Di). (2019). *EEG Signal Processing and Feature Extraction*. Springer. https://doi.org/10.1007/978-981-13-9113-2

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, *13*(4), 411–430. https://doi.org/10.1016/S0893-6080(00)00026-5

Jolicœur, P., Sessa, P., Dell'Acqua, R., & Robitaille, N. (2006). On the control of visual spatial attention: Evidence from human electrophysiology. *Psychological Research*, *70*(6), 414–424. https://doi.org/10.1007/s00426-005-0008-4

Kang, M.-S., & Woodman, G. F. (2014). The neurophysiological index of visual working memory maintenance is not due to load dependent eye movements. *Neuropsychologia*, *56*, 63–72. https://doi.org/10.1016/j.neuropsychologia.2013.12.028

Kowler, E. (1995). Eye movements. In *Visual cognition: An invitation to cognitive science, Vol. 2, 2nd ed* (pp. 215–265). The MIT Press.

Lagerlund, T. D., Sharbrough, F. W., & Busacker, N. E. (1997). Spatial Filtering of Multichannel Electroencephalographic Recordings Through Principal Component Analysis by Singular Value Decomposition. *Journal of Clinical Neurophysiology*, *14*(1), 73.

Lawrence, M. A., & Lawrence, M. M. A. (2016). Package 'ez'. *R package version*, *4*(0).

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, second edition*. MIT Press.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience & Biobehavioral Reviews*, *62*, 100–108. https://doi.org/10.1016/j.neubiorev.2016.01.003

Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2010). Visual Short-term Memory Capacity for Simple and Complex Objects. *Journal of Cognitive Neuroscience*, *22*(3), 496–512. https://doi.org/10.1162/jocn.2009.21214

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*(3), Articolo 3. https://doi.org/10.1038/nrn1348

Meconi, F., Doro, M., Schiano Lomoriello, A., Mastrella, G., & Sessa, P. (2018). Neural measures of the role of affective prosody in empathy for pain. *Scientific Reports*, *8*(1), Articolo 1. https://doi.org/10.1038/s41598-017-18552-y

Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology*, *47*(6), 1142–1150. https://doi.org/10.1111/j.1469-8986.2010.01015.x

Nicenboim, B. (Regista). (2020, marzo 27). *eeguana: A package for flexible manipulation of EEG data*. https://doi.org/10.5281/ZENODO.3730553

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Plöchl, M., Ossandón, J., & König, P. (2012). Combining EEG and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, *6*. https://www.frontiersin.org/articles/10.3389/fnhum.2012.00278

Pontifex, M. B., Gwizdala, K. L., Parks, A. C., Billinger, M., & Brunner, C. (2017). Variability of ICA decomposition may impact EEG signals when used to remove eyeblink artifacts. *Psychophysiology*, *54*(3), 386–398. https://doi.org/10.1111/psyp.12804

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330. https://doi.org/10.3758/s13423-011-0055-3

Roy, Y., & Faubert, J. (2023). Is the Contralateral Delay Activity (CDA) a robust neural correlate for Visual Working Memory (VWM) tasks? A reproducibility study. *Psychophysiology*, *60*(2), e14180. https://doi.org/10.1111/psyp.14180

Sazgar, M., & Young, M. G. (2019). EEG Artifacts. In M. Sazgar & M. G. Young (A c. Di), *Absolute Epilepsy and EEG Rotation Review: Essentials for Trainees* (pp. 149–162). Springer International Publishing. https://doi.org/10.1007/978-3-030-03511-2_8

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, *118*(1), 98–104. https://doi.org/10.1016/j.clinph.2006.09.003

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

Sessa, P., Luria, R., Gotler, A., Jolicœur, P., & Dell'acqua, R. (2011). Interhemispheric ERP asymmetries over inferior parietal cortex reveal differential visual working memory maintenance for fearful versus neutral facial identities. *Psychophysiology*, *48*(2), 187–197. https://doi.org/10.1111/j.1469-8986.2010.01046.x

Sheppard, A. L., & Wolffsohn, J. S. (2018). Digital eye strain: Prevalence, measurement and amelioration. *BMJ Open Ophthalmology*, *3*(1), e000146. https://doi.org/10.1136/bmjophth-2018-000146

Stone, J. V. (2002). Independent component analysis: An introduction. *Trends in Cognitive Sciences*, *6*(2), 59–64. https://doi.org/10.1016/S1364-6613(00)01813-1

Sugimoto, H., Ishii, N., Iwata, A., Suzumura, N., & Tomita, T. (1978). On the stationarity and normality of the electroencephalographic data during sleep stages. *Computer Programs in Biomedicine*, *8*(3–4), 224–234. https://doi.org/10.1016/0010-468x(78)90029-6

Sussman, E. S. (2007). A New View on the MMN and Attention Debate. *Journal of Psychophysiology*, *21*(3–4), 164–175. https://doi.org/10.1027/0269-8803.21.34.164

Teplan, M. (2002). FUNDAMENTALS OF EEG MEASUREMENT. *MEASUREMENT SCIENCE REVIEW*, *2*.

Verleger, R. (1991). The instruction to refrain from blinking affects auditory P3 and N1 amplitudes. *Electroencephalography and Clinical Neurophysiology*, *78*(3), 240–251. https://doi.org/10.1016/0013-4694(91)90039-7

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), Articolo 6984. https://doi.org/10.1038/nature02447

Zou, Y., Nathan, V., & Jafari, R. (2016). Automatic Identification of Artifact-Related Independent Components for Artifact Removal in EEG Recordings. *IEEE Journal of Biomedical and Health Informatics*, *20*(1), 73–81. https://doi.org/10.1109/JBHI.2014.2370646