UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Psicologia Generale
Dipartimento di Psicologia dello Sviluppo e della Socializzazione

Corso di laurea magistrale in Neuroscienze e Riabilitazione Neuropsicologica

TESI DI LAUREA MAGISTRALE

Investigating the relationship between visual working memory and consciousness: Insights from a study combining binocular rivalry and EEG

Relatrice
Prof.ssa Paola Sessa

Laureanda: Nicole Genero

Matricola: 2086432

Anno Accademico 2023/2024

TABLE OF CONTENTS

ABSTRACT

In recent decades, a significant debate has emerged in literature regarding the nature of the representations that working memory can maintain. Specifically, some researchers have suggested the possibility that working memory may be capable of holding not only conscious, but also unconscious information.

It is now well established that the study of working memory is of considerable importance, as it correlates with general cognitive abilities. Although numerous studies have investigated how resources are allocated within working memory, the model that has received the most empirical support posits that the number of possible representations is discrete. Furthermore, the literature has shown great interest in Contralateral Delay Activity (CDA), an event-related potential considered to be an index of working memory capacity, as its amplitude increases linearly with the number of representations that are being held in working memory. As for the nature of these representations, it is essential to underline that consciousness represents one of the most debated topics in neuroscience. Specifically, modern neuroscience has developed several theories attempting to define consciousness or identify its neural correlates, often yielding contradictory results. Currently, since pinpointing its precise neural basis seems unattainable, research has shifted towards explaining the phenomena associated with consciousness. In this regard, Dehaene and Naccache's Global Neuronal Workspace Theory (2001) supports the idea that certain cognitive processes may occur without conscious awareness, due to the interaction of interconnected brain regions. The present study, as described in this thesis, builds upon this principle to explore the possibility that working memory can also maintain information of which the individual is not aware. Specifically, the study analyzes the Contralateral Delay Activity using a combination of electroencephalography and a binocular rivalry paradigm. In details, participants are asked to observe pairs of Gabor patches on a monitor through a stereoscopic mirror, creating a condition of binocular rivalry, and to indicate via button press the orientation of the bars within the observed stimuli. The results support the existence of unconscious processing of visual stimuli and highlight how working memory can filter out irrelevant representations to retain only the pertinent information.

INTRODUCTION

This thesis originates from the pre-graduate internship conducted at the electroencephalography laboratory at the Department of Developmental Psychology and Socialisation at the University of Padua. Specifically, this research represents a collaboration between the University of Padua and Tel Aviv University, particularly with Professor Roy Luria and his research team in Israel.

This research aims to investigate the relationship between working memory, specifically visual working memory, and consciousness. The experimental procedure seeks to clarify the nature of the representations that are maintained in working memory as a person interacts with the external world. Furthermore, the study aims to determine whether these representations can have an unconscious nature. While literature provides some evidence supporting this hypothesis, no findings can be deemed conclusive due to difficulties in measuring unconscious phenomena.

The first theoretical chapter of this thesis revisits the primary descriptive models of working memory, focusing specifically on slot-based models, which posit that working memory has a limited capacity. These models are supported by substantial empirical evidence, particularly the contralateral delay activity, an event-related potential whose amplitude increases linearly with the number of items retained in working memory until it reaches an asymptote at maximum capacity. The Contralateral Delay Activity is a valuable wave for studying the processing of stimuli, as observing its amplitude allows researchers to determine whether a stimulus is actually present in working memory. Consequently, the experimental paradigm employed in this study examines Contralateral Delay Activity amplitude during stimulus observation on a monitor.

The second chapter provides a critical overview of the most influential theories of consciousness to date. More specifically, there are several consciousness theorizations and each theory addresses different aspects of the conscious experience. As highlighted in literature, cognitive neurosciences often focus on identifying the neural correlates of consciousness. However, this approach poses challenges in obtaining reliable results due to the difficulty in measuring unconscious processes with current psychophysiological recording techniques and experimental methodologies. Nonetheless, a particularly noteworthy theory is the Global Neuronal Workspace Theory proposed by Dehaene and

Naccache in 2001, rooted in Baars' 1998 Global Workspace Theory. Dehaene and Naccache's theory posits the existence of a distributed neural network that facilitates communication between cognitive systems, enabling consciousness. Additionally, this theory aligns with our hypothesis by suggesting the existence of unconscious cognitive processing. The second chapter also introduces binocular rivalry as a technique for studying unconscious processes without requiring participants to report their level of awareness during the experiment. This method reduces overall noise coming from the responses given by participants, therefore providing a more reliable measure of unconscious processes. Given these advantages, we incorporated binocular rivalry into our experimental paradigm.

The third chapter provides a detailed description of the experimental methodology that has been adopted. Briefly, the study involves a computer-based task lasting approximately 32 minutes, excluding breaks, during which subjects' electroencephalographic activity is recorded. This recording allows for the calculation of Contralateral Delay Activity amplitude variations during the viewing of stimuli that are presented via a stereoscopic mirror, to create a binocular rivalry condition. Thus, two different images are presented to the retinas, but the participant is aware of only one image. The participant indicates which image they perceive (A or B) by pressing the corresponding keys on the keyboard.

In conclusion, the analyzed data appear to support the initial hypothesis that working memory retains even unconscious information. Nevertheless, the limited sample size and experimental design could be further refined to enhance research conditions, thereby achieving more reliable results. However, the methodology employed in this study still provides excellent basis for future research in this field, considering its numerous advantages and relatively simple applicability.

CHAPTER 1

1.1 *VISUAL WORKING MEMORY*

1.1.1 Rise and definition

For many decades, the large-scale theories of cognition have been considering working memory (WM) to be a core element in cognitive processes. In the 1960s researchers strongly agreed on the fact that the human memory system could be divided in two components. More precisely, they theorized the presence of a short-term store that could hold a limited amount of information for a few seconds, namely the short-term memory (STM). Furthermore, a long-term memory system was thought to allow the storage of vast amounts of information for longer periods of time.

Baddeley and Hitch (1974) argued that the short-term store could also function as a working memory, meaning that the STM's role was not just storing the information gathered through attentive processes to store them in the long-term system. In fact, according to the authors, the STM could also serve to hold information online to manipulate them, in order to complete cognitive operations. More precisely, it was observed that performance in cognitive tasks could be significantly disrupted by loading the STM to capacity, meaning that the time required to complete the task increased with load. Subsequently, these authors theorized a model which could account for the aforementioned experimental data and many others.

According to Baddeley's multicomponent model (1974), WM can be split in several components, and it is not considered a unified construct anymore. There were three main components described in the original model: the central executive, the phonological loop and the visuo-spatial sketchpad. While the phonological loop allows manipulating and storing verbal data, the visuo-spatial sketchpad selectively gathers visuo-spatial information. Then, the central executive acts as an attentional controller and coordinates the data, aided by the other two subsystems. Baddeley (2000) added a fourth system, the episodic buffer, to explain how the other three systems operate together to interface with long-term memory. The author states that the episodic buffer comprises a limited capacity unit that allows information to be stored temporarily. This information, which are collected from the other subsidiary systems and from long-term memory, are then bound in a unitary episodic representation (Baddeley, 2000), due to the action of the episodic buffer. In addition, this model assumes conscious awareness to be the main retrieval mode from the buffer.

It is fundamental to note that Baddeley's model (1974; 2000) is not the only relevant theorization described in scientific literature regarding working memory. Atkinson & Shiffrin (1968) proposed a theoretical model which divides memory into three structural systems: the sensory register, the short-term store, and the long-term store. The short-term store is essentially the working memory and gathers inputs from the other two components. The multicomponent model of working memory (Baddeley & Hitch, 1974) describes a three-part working memory model that was formulated as an alternative to Atkinson's short-term system.

In the last decades, several studies have confirmed that phonological and visuospatial processing require separate mechanisms and involve different neural substrates.

Originally, verbal tasks were more frequently used to assess the cognitive dynamics of working memory (Luck & Vogel, 2013). An example is provided by the digit span tasks, which consists in the repetition of a series of numbers in the same order. Another example is constituted by the complex span task, which requires the participant to alternate between two different tasks that exacts both processing and storing information in memory (e.g., solving mathematical operations while remembering unrelated words, as in the operation span task used by Unsworth et al., 2005). However, the past 25 years have witnessed a rise of research on visual working memory.

Visual working memory (VWM), or visual short-term memory (VSTM) can be defined as an online limited workspace that allows to actively maintain visual information for a few seconds. Consequently, this information can be manipulated by higher order cognitive functions to complete an ongoing task (Luck & Vogel, 2013).

In compliance with Luck and Vogel (2013), qualifying as VWM implies meeting three requirements. First, the fact that the representation of the information is visual in nature is necessary. So, acquiring that information through the visual modality is still necessary but not sufficient. Second, what distinguishes VWM from longer-term memories is that VWM is based on active maintenance. This specifically means that a VWM representation is preserved through continuous, energy-consuming neural activity rather than by altering synaptic strength. Third, the visual representations need to be employed to support broader cognitive tasks, since VWM contains the term “working” by definition.

1.1.2 Experimental paradigms that measure visual working memory

The most common experimental procedure that has been used to study visual working memory involves a short presentation of a set of stimuli, followed by a delay period and a memory test regarding the previously view items. One of the most common procedures, that has been highly modified since it was firstly utilized, is the Change Detection Task (Philips, 1974). This experimental paradigm was firstly introduced by Philips in 1974. More precisely, Philips (1974) created a pattern of colored squares that randomly filled a square matrix and presented this pattern for 1 s to the observers. A delay period of different intervals was introduced, and then another pattern followed. This last pattern was either identical or similar to the pattern showed before the delay period, and the participant had to declare if the last-viewed pattern was the same or differed from the original one (*Figure 1*). Philips (1974) observed that the performance could be less accurate if the pattern was high in complexity, of if the delay period was quite long.

In the last few decades, several variants of the Change Detection Task have been developed, and some of them will be discussed below. Notably, this task uses a sequential comparison procedure. However, there are other ways to study VWM, even though this paradigm remains the most frequently used.

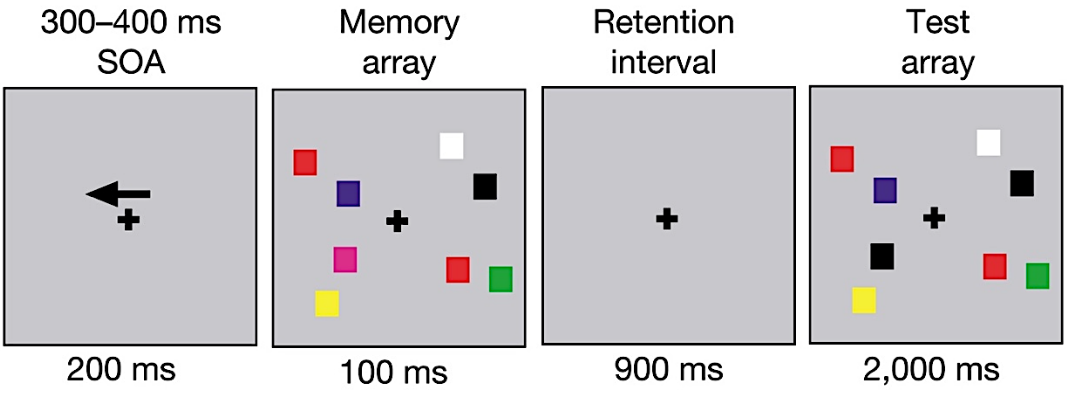


Figure 1. An example of the Change Detection Task, from Vogel and Machizawa (2004).

1.1.3 VWM capacity

As stated above, VWM has a limited capacity. Research about VWM capacity has been going on since the 1960s. More precisely, Sperling (1960) determined that each participant could report only a limited amount of symbols. Specifically, participants were exposed to both letters and number from 15 to 500 msec in seven different experimental conditions. The results indicated that the average of reported symbols was slightly over four letters, and that stimuli with four or fewer letters were reported correctly approximately 100% of the time (Sperling, 1960). However, the most substantial limitation of this study is that it is not clear whether the to-be-remembered items were stored in visual or phonological representations, considering them being both letters and numbers.

There are two main categories of theories about VWM capacity that have been proposed and discussed during the last few decades: slot-based theories (or discrete slots classes) and resource-based theories (or continuous resources classes).

Resource-based theories propose that visual working memory capacity is a flexible resource that can be allocated across all items in a display. However, as the number of items increases, fewer resources are available to be allocated to each item, leading to reduced precision. Consequently, there will be an augmentation of neural noise when the set size grows (Luck & Vogel, 2013). In compliance with the resource-based theory, VWM resources can be distributed flexibly to create either a few high-detail representations or many low-detail representations (Zhang & Luck, 2008).

According to slot-based theories, the number of items that can be stored in VWM is limited, and its maximum value is defined as K_{max} (Luck & Vogel, 2013). Subsequently, if the number of to-be-remembered items exceeds K_{max}, the information that exceed that value will be lost. Note, however, that there is a moderate variability of K_{max} value among different trials, so K_{max} also varies according to task difficulty, other than among individuals. Therefore, object complexity can play a role in defining *Kmax*.

Many empirical confirmations for both theories have been described in literature. Bays and Husain (2008) provided evidence in favor of the resource-based model for VWM capacity, testing whether VWM is a finite resource that is dynamically and flexibly distributed among all items in the visual field. Specifically, the authors used a variation of the spatial memory paradigm. Participants were asked to recall the location and orientation of multiple visual items (e.g., a sample array of colored squares) after a delay by a probe stimulus. In other words, the subjects were asked to determine whether the probe stimulus was displaced to the right or to the left. In accordance with resource-based theories, the performance was almost flawless, even when the displacement was sufficiently large and the set size increased. Furthermore, the accuracy of the reports steadily decreased as the set size increased. This pattern strongly supports continuous resources classes, as slot-based models would predict that participants should not make any errors until K_{max} is reached. At first, this could indicate unequivocable evidence of resourced-based theories. However, the obtained results (Bays and Husain, 2009) seem to reflect the usage of a guessing strategy (Thiele et al., 2011). In fact, Thiele and colleagues claim that the methodological design used by Bays and Husain (2009) was flawed, because participants could predict the correct solution due to the test array itself and without using working memory processes. Specifically, the authors found a correlation between the position of the item at test and the direction of displacement. Some trials, in fact, had an obvious answer because the test item was on the right or left edge. For example, if the item was on the right edge, the item itself could not have been displaced to the left, but only to the right side. Hence, Thiele and colleagues (2011) modified the paradigm by presenting the probe at the horizontal centre on the screen. Results indicate that performance in the test-at-center condition is perfect, or almost perfect, only when the set size ranges from one to three items with large displacement. When the set size increased to eight, the results obtained by Bays and Husain (2009) were not replicated.

By contrast, evidence of slot-based theories is widely described in literature.

Whilst James in the 1890s assumed an unlimited capacity, Miller's (1956) theorization "magical number seven, plus or minus two" probably remains the most relevant paper investigating the short-term memory (STM) storage capacity (Cowan, 2001). According to Cowan (2001) the number of discrete slots ranges from three to four items in the adult population. More precisely, Cowan (2001) refers to "chunks", which are small groups of information that are bound together and serve the purpose of bypassing the capacity of working memory. Luck and Vogel (1997) demonstrated that it is possible to remember both colors and orientations of four objects, which suggests that visual working memory holds representations of complete, integrated objects and not only of individual features, such as color. In details, the experimental procedure applied a variant of the change detection task created by Philips (1974) that consists of viewing two different sequential arrays, one sample and one test, and deciding whether they differ from one another. The first experimental set (Luck & Vogel, 1997) tested VWM capacity for simple colors and results indicated a nearly perfect performance for arrays of 1-3 items. Accuracy decreased linearly when set size grew from 4 stimuli to 12. The authors ensured that the observed changes in accuracy were caused by limitations in storage capacity rather than differences among low-level visual processes. The authors then assessed if visual working memory held representations of integrated objects rather than just individual features. Observers executed the aforementioned task (e.g., Philips, 1974) with arrays of 2, 4, 6 colored bars of different orientations. Color, orientation or both could vary among trials. Results indicate that the performance was almost identical when subjects had to hold both color and orientation (e.g., conjunction condition). According to this finding, authors conclude that VWM can store integrated percepts, as the verbal working memory can store chunks (Miller, 1956).

During more recent years, evidence of discrete slots also derives from continuous report experiments, which allow to establish whether VWM representations become less accurate as set size grows (Luck & Vogel, 2013). Zhang and Luck (2008) aimed at providing evidence of discrete slots by utilizing a short-term recall paradigm, in which each trial is constituted by a sample array, a retention interval and then the test display. The aim of this experiment is to report the color of a remembered item (e.g., sample array) by clicking on a color wheel (e.g., test display). If the probed item is present in WM, the observer should be able to report a color, in the color wheel, which is similar to the original one. Vice versa, according to slot-based theories, the error distribution should be random if the participant has no information about the color, since the cued item is not present in memory. According to the results, when given more than a few simple stimuli, humans keep a high-detailed memory of only some of them and do not remember any details about the rest (Zhang & Luck, 2008). In the same paper, Zhang and Luck also introduced the possibility that WM capacity could be more accurately described by either "slots + resources" or "slots + averaging" models. The former assumes that most of the resources are allocated to a single representation, increasing its precision, whilst leaving less resources for the other representations, which would then be highly imprecise.

Instead, the latter model theorizes that if three memory slots are available, then they could all be utilized to represent a single stimulus. Thus, if each representations holds an independent sample of the stimulus, and observers report the average of three such representations during the test, the precision of the report will improve. However, the authors demonstrated that the usage of a slot + resources model would be highly unlikely. By contrast, the authors highlight that a model constituted by a small set of discrete, fixed-resolution representations can yield a quantitative account of memory performance among many experimental paradigms. Moreover, the slots + averaging model perfectly describes the definition of resource and outlines the limits on how resources can be distributed among different stimuli (Zhang & Luck, 2008).

Several other evidence in favor of slot-based models have been described. For example, Zhang and Luck (2011) examined whether participants could trade quantity for quality in working memory when given monetary incentives to do so. In other words, the authors hypothesized that increasing the quantity of items held in VWM beyond K_{max} could decrease the details of the representations. However, the observers were not capable of increasing K_{max}, even when given incentives.

Furthermore, evidence of discrete slots in working memory capacity are not only provided by behavioral data. An unequivocable proof of a limited number of slots is given by ERP and fMRI measures of VWM delay activity, which increases as the set size grow bigger up until K_{max}, when it reaches an asymptote (Luck & Vogel, 2013).

1.1.4 Individual and group differences in visual WM capacity

WM capacity is thought to explain several individual differences in cognitive abilities. Moreover, according to Johnson et al., (2013) WM is significantly correlated with a measure of higher cognitive functions (e.g., the T score in the MATRICS Consensus Cognitive Battery).

Nevertheless, WM capacity is impaired in many psychopathological conditions and can significantly differ among different populations (Luck & Vogel, 2013). For example, individuals with a schizophrenia diagnosis have been proven to display reduced VWM capacity (i.e., K_{max}) compared to healthy controls (Gold et al., 2013). More precisely, Gold and colleagues modeled the change detection task after the task used by

Luck and Vogel (1997), which is described above. Each subject, both healthy and suffering from schizophrenia, was tested with set sizes that ranged from 2, 3 4 and 6 bars and there were three experimental conditions: color, orientation, conjunction of color and orientation. These three conditions refer to the item's characteristic that changed between the sample array and the test trial. Results illustrate how schizophrenia patients have clear and demonstrable deficits in the quantity of information they can store in WM. Moreover, the value K, that indicates the average number of items held in WM, decreases linearly when the set size increases from 4 to 6. This latter result can be observed only for individual with schizophrenia and not for healthy participants. Considering that VWM is correlated with broader cognitive ability, it can be inferred that some cognitive deficits could be reduced by a specific treatment that targets, and consequently normalizes, the patient's VWM capacity.

In addition, an experiment carried out by Vogel et al., (2005) suggests allocating attention to irrelevant information to be the main factor to cause observable differences in overall storage capacity among healthy individuals. Thus, results indicate that selection efficiency varies significantly among different people, and it is also predicted by individual WM capacity.

It is fundamental to note that differences in filtering ability are not able to explain all group differences in VWM capacity that can be found in literature. For example, a study by Lee et al., (2010) comparing the performance of healthy controls and medication-withdrawn patients with Parkinson's Disease highlighted how patients were both impaired when they had to filter out distractors, and they could also maintain fewer stimuli in WM than controls (Lee et al., 2010). This finding is also consistent with the neuroanatomical aspects of Parkinson's Disease. In fact, this disease highly affects the frontostriatal circuit, which supports WM processes. Moreover, this finding also supports the theorization which states that basal ganglia can exert attentional control over access to WM (McNab and Klingberg, 2008).

Other examples are provided by additional studies involving patients with schizophrenia (Hahn et al., 2010; Mayer et al., 2012). In compliance with these studies, the patient's difficulties in VWM tasks may be caused by a tendency to hyperfocus their attentional resources on a reduced number of items. Specifically, individuals who suffer from schizophrenia tend to hyperfocus on visual information that are perceptually relevant. Mayer and colleagues (2012), in fact, proved that schizophrenia patients show compromised filtering ability and slow disengagement when they were shown perceptually salient distractors.

Nevertheless, investigating individual differences in VWM processes seems to be fundamental to better understand certain diseases and, to a certain extent, to develop targeted treatments to ameliorate patients' cognitive abilities and their overall quality of life (Luck and Vogel, 2013).

1.1.5 Neural bases and mechanisms of VWM

The majority of neural network models propose that representations in VWM are preserved through recurrent feedback loops, where information circulates from one group of neurons to another and then returns to the original group (Zhang and Luck, 2013). A recurrent feedback loop can effectively maintain a single item, but it becomes challenging to preserve multiple distinct items without their representations merging into one. To address this issue, all neurons representing a specific item are connected in a synchronously firing cell assembly, with only one such assembly active at any given moment. A cell assembly corresponds to a neuronal group that codes different stimuli in VWM. Synchrony within a cell assembly helps sustain recurrent activation, while asynchrony between different assemblies prevents interference between the representations of various items. A synchronous cell assembly efficiently corresponds to a slot in VWM (Zhang and luck, 2013).

According to these models, the role of neuronal firing seems fundamental. Specifically, cell assemblies go off after firing and if too much time passes before firing again, then the visual representation is lost. Subsequently, to hold representations of multiple stimuli, it is necessary both to prevent too many cell assemblies from firing simultaneously and to avoid the decay of a cell assembly (Raffone and Wolters, 2001). The model suggested by Raffone and Wolters (2001) was empirically tested through electrophysiological recordings and it accounts for the results found by Luck and Vogel (1997). The latter found that the observer integrates multiple characteristics of an object (e.g., color and orientation) to hold just one representation of the stimuli in visual working memory. Thus, the fact the neurons that code for different characteristics can synchronize into one cell assembly to form a visual representation seem quite consistent with the findings obtained from Luck and Vogel (1997).

Furthermore, most models, including the theorization from Raffone and Wolters (2001), assume the involvement of only two cortical areas in WM processes. These areas are the inferotemporal (IT) and prefrontal (PFC) cortices.

It has been widely demonstrated that the PFC plays a major role in every aspect of WM. As proven by Baddeley (1986, 1992) individuals who suffer from frontal lobe damage, qualifying for a diagnosis of "dysexecutive syndrome", have WM disruption.

According to Ungerleider and colleagues (1998) VWM involves a distributed neural network that includes both posterior areas in the visual cortex and prefrontal cortex. Moreover, within the visual cortex, the dorsal stream is involved in WM for spatial locations, while the ventral stream is involved in WM for visual objects. Ungerleider (1998) studied both monkeys and humans observing that, in monkeys, the domain specificity that could be observed in the distinction between ventral and dorsal pathway could also be observed across the prefrontal cortex. More precisely, ventral prefrontal regions support WM processes for objects and dorsolateral regions support WM for spatial locations. In humans, it is widely known that the inferior temporal cortex supports object vision in the ventral stream while the dorsal stream has an upper location in parietal cortices.

While the areas supporting working memory processes in the ventrolateral prefrontal cortex are similarly located in monkeys and humans, the areas adhibited for spatial WM processing are more posterior and located in the dorsal prefrontal cortex in humans.

1.2 ***CONTRALATERAL DELAY ACTIVITY (CDA) AS AN INDEX OF VISUAL WORKING MEMORY***

1.2.1 Brief history of the contralateral delay activity

In recent years, there has been a growing interesting in studying visual working memory with electrophysiological techniques. As mentioned above, VWM has been proven to have a set capacity even with neuroimaging techniques. Specifically, the contralateral delay activity (CDA) is a negative slow wave whose amplitude is positively correlated with the number of items that are being held in working memory (Luria et al., 2016).

Ruchkin and colleagues (1990) were the first to observe a sustained EEG activity during the retention interval in a WM task. More precisely, the authors recorded the brain's activity while participants were asked to memorize a stimulus while varying the number of items (1, 3 or 6) in the task stimulus (Ruchkin et al., 1990). Event-related potentials (ERP) activity was recorded for 2450 msec after task stimulus offset. The authors were able to observe two different waves: a positive posterior wave around Pz and a negative frontal wave. The posterior positive wave was clearly associated with retention and increased with memory load during the memory task. Moreover, this component was not affected by performance, so it was thought not to be related to accuracy. However, it is important to note that at the time ERP analysis was not as precise as it is today. In fact, according to Luria (2016) it is possible to assume that the observed posterior wave could also be due to several different processes occurring during the task, that are non-mnemonic processes (e.g., arousal or attentional processes).

A few years later, Klaver and colleagues (1999) were able to find a negative activity over posterior electrodes again, but this result was overlooked due to poor methodological design. At the time, this wave was thought to be caused by limited resource processing (Luria et al., 2016).

In 2004 Vogel and Machizawa (2004) developed a bilateral version of the change detection paradigm. Having a bilateral version served the purpose of observing a highly specific mnemonic activity by calculating the difference between the addended side of the screen and the non-attended side of the screen. Consequently, the authors were able to observe the CDA as a negative slow wave that persisted through all the retention interval and, as Ruchkin (1990) had found many years prior, whose amplitude was sensitive to the number of items held in memory. As mentioned above while illustrating individual differences on VWM capacity, the CDA had also been used by Vogel (2005) to demonstrate that, for some individuals, having reduced VWM capacity can be caused by an impairment in filtering out irrelevant information, therefore saturating the VWM limited workspace (Vogel et al., 2005).

Firstly, the CDA was observed during change detection task paradigms, hence the name “contralateral delay”. However, later studies found a similar activity in working memory task, but not during the retention interval. In compliance with those findings, some authors labeled this posterior activity with several different names. Notably, different studies referred to VWM activity as Contralateral Negative Slow Wave (CNSW; Klaver et al., 1999), Sustained Posterior Contralateral Negativity (SPCN; Brisson and Jolicoeur, 2007; Perron et al., 2009) and Contralateral Search Activity (CSA; Emrich et al., 2009).

1.2.2 CDA as a difference wave

The CDA seems to be a robust neural correlate of VWM according to Roy and Faubert (2022). In fact, the authors were able to reproduce findings from very diverse experiments, resulting in a clear extraction of the CDA signal (Roy and Faubert, 2022).

According to Vogel and Machizawa (2004) the CDA component was elicited 200 ms after the onset of the memory array at OL/OR sites, which correspond to PO7/PO8 according to the extended 10-20 system. Moreover, this slow wave persisted throughout the retention interval (Balaban and Luria, 2019).

More precisely, the CDA is time-locked to the onset of the memory array, and it is measured during the retention period, before the appearance of the test array (Luria et al., 2016).

The CDA is calculated by subtracting the amplitude between the contralateral side and the ipsilateral side of fixation (*Figure 2*). In fact, low level and early perceptual processing are reflected in the amplitude on the ipsilateral side, while the amplitude on the contralateral side is related to both low level processes and VWM activity (Luria et al., 2016).

By subtracting the two waves it is possible to get a clean signal, which does not include neither local noise nor low level visual processing. Notably, it is not possible to know whether the two pre-subtracted waveforms are positive or negative, because the only information provided by the CDA is that the contralateral wave is more negative than the other.

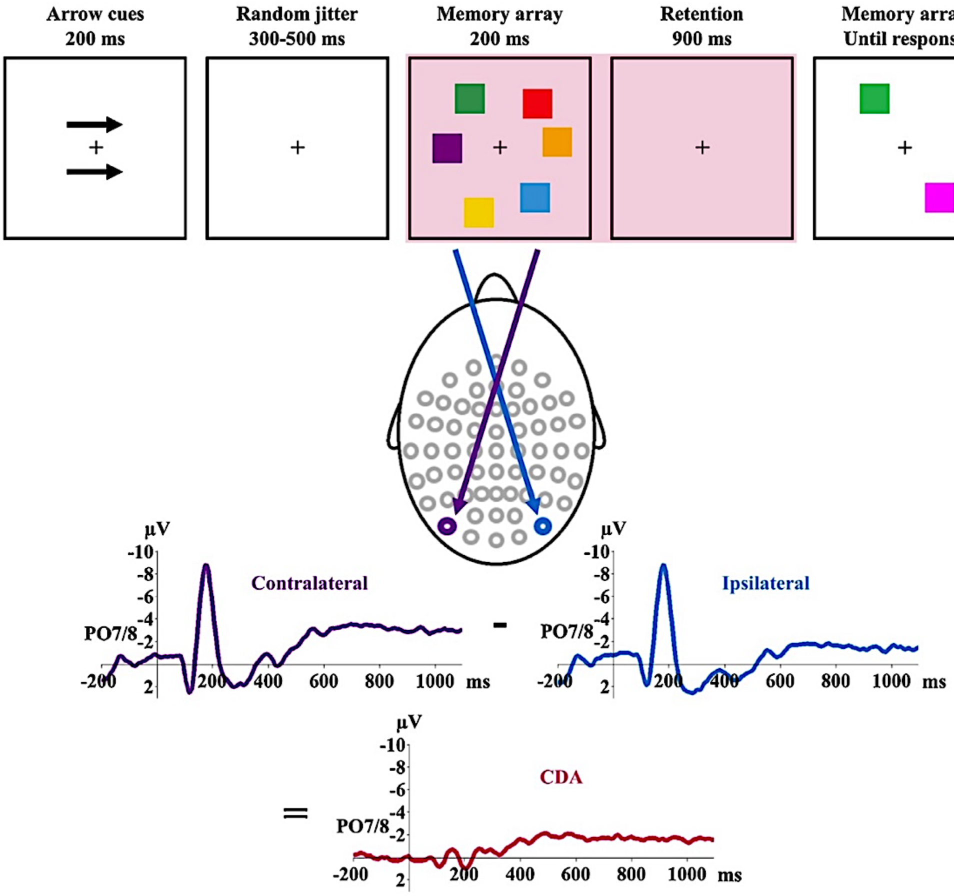


Figure 2. An illustration of a typical change detection trial and the resulting CDA waveforms from Luria and colleagues (2016). The CDA results from the subtraction of ipsilateral activity from contralateral activity, and it is measured throughout the retention interval.

Arend and Zimmer (2011) investigated whether the amplitude of the contralateral delay activity (CDA) could be influenced by the number of objects present in the opposite hemifield. Their study established two conditions, including one where the number of items on one side differed from the number on the contralateral side. The results showed that the contralateral wave was affected by the number of items in the ipsilateral side only when a single item was held in VWM on the contralateral side. Conversely, similar findings were observed for the ipsilateral wave. These results suggest that the processing of irrelevant items depends on the VWM load (Arend and Zimmer, 2011).

In addition, Liesefeld and colleagues (2014) found a correlation between measures of filtering in a change detection task and the contralateral amplitude. However, the authors could not clarify whether the CDA, considered as a difference wave, had worse or better correlations (Liesefeld et al., 2014). Subsequently, the difference between the two waves should be investigated to determine their specific underlying cognitive mechanisms.

1.2.3 CDA as an index of visual working memory capacity

As it was emphasized above, the fundamental quality of the CDA is that its amplitude rises according to the number of items that are being held in VWM (Luria et al., 2016). Notably, this amplitude increase is not due to an augmented task complexity, since the CDA reaches an asymptote at 3 to 4 objects. Thus, CDA reaches an asymptote when the maximum capacity of VWM, or K_{max}, is reached. Moreover, the CDA asymptote itself is correlated with individual VWM capacity.

Luria and colleagues (2016) carefully examined 11 studies to better quantify the aforementioned relation between VWM load and CDA. The meta-analysis revealed a combined correlation of $r= 0.596$ and a 95% confidence interval between 0.510 and 0.670 (Luria et al., 2016). This finding provides compelling evidence that the CDA is responsive to the number of objects held in VWM. Moreover, results suggest both that the CDA amplitude increases with a rise in set size, and that this increase is less pronounced in individuals with lower VWM capacity.

There are many other studies that have validated the CDA to be an index of VWM capacity. For example, McCollough and colleagues (2007) succeeded in demonstrating that the CDA amplitude is smaller on incorrect response trials compared to correct trials. This finding could imply that this component may be fundamental to answer correctly on a given trial (McCollough et al., 2007), and it is also consistent with the assumption that mistakes occur because relevant information can be lost in WM (Luria et al., 2016).

Notably, in accordance with several studies (Gao et al., 2013; Ikkai et al., 2010; Ye et al., 2014) CDA amplitude and low accuracy performance dissociate. Specifically, no significant effects were found when the stimulus contrast was manipulated in the change detection paradigm.

Ikkai and colleagues (2010) note that there are two alternative accounts that could cause to wrongly assume that the CDA is responsive to VWM load. First, CDA amplitude may be related to the increasing perceptual demands on the screen. Second, the CDA could be modulated by the number of locations on the display towards which attention is allocated, since all previous study confounded the number of to-be-remembered items with the number of possible locations (Ikkai et al., 2010). The authors' experimental designs were specifically made to clarify the aforementioned CDA's characteristics. Results indicate that increasing the perceptual demands of the task, by significantly lowering contrast, did not lead to an increase in CDA amplitude in the first experiment. Moreover, the second experiment showed CDA amplitude to be decoupled from location numbers (Ikkai et al., 2010).

To further prove that CDA is responsive to the number of items that are being held in VWM, neuroanatomical evidence should also be considered. More precisely, Todd and Marois (2005) found evidence that Intraparietal Sulcus (IPS) is involved in VWM storage capacity, which is an index of VWM. More specifically, the authors estimated that IPS accounts for about 40% of variance in VWM load among subjects. This is consistent with the established finding that Posterior Parietal Cortex (PPC) is the locus for information storage in VWM (Todd and Marois, 2005). A few years before, Jovicich (2001) was also able to demonstrate that the IPS activity can be modulated by varying the number of stimuli that are being tracked by the observer (Jovicich, 2001).

1.2.4 Using the CDA to study visual working memory

The main advantage of utilizing ERPs to study cognitive processes relies on the capacity to unveil the subprocesses that undergo VWM, which would otherwise be impossible to observe. This is caused by the fact that behavioral data can only show the output result of the process, which always comprise different processes unrelated to VWM. As a consequence, since Vogel and Machizawa (2004) firstly introduced the CDA as a marker of VWM, there has been a growing interest in using the former as a dependent measure in several tasks that are typically used to study WM.

Notably, the more traditional paradigms that are frequently used to study VWM (such as change detection task) involve stimuli that have to be maintained in memory due to their subsequent disappearance during the delay period. However, Tsubomi and colleagues (2013) successfully proved that the same neural activity that is found for items that are no longer visible can be found when processing items that remain continuously visible. Together, the findings of this study appear to be inconsistent with the traditional and historical theorization, which states that VWM limitations can be observed when the participant has to store visual inputs that have to be remembered due to their disappearance. Therefore, the authors conclude that taken together, the CDA and VWM set capacity are not exclusive to the maintenance of sensory inputs that are no longer present in the visual scene, but they also modulate conscious representations of input that do not disappear and can be seen (Tsubomi et al., 2013). In other words, these findings suggest that WM is more about "working" (i.e., online processes) rather than "memory" (Luria et al., 2016).

In literature experimental paradigms that involve the CDA analysis usually fall into two main categories. Some studies track CDA raw amplitude, but most studies observe the CDA difference between two different experimental conditions. The latter is more frequent, since having two experimental conditions reduces general noise by allowing to control for several other components. As for the first category of studies, in some cases the CDA raw amplitude was found to be correlated with VWM capacity of the participant (Luria and Vogel, 2011b; Voytek and Knight, 2010; Wiegand et al., 2013) and in other cases it correlated with accuracy or reaction times (RTs) (Carlisle et al., 2011; Gunseli et al., 2014; Woodman and Arita, 2011).

The majority of empirical evidence about CDA comes from studies that investigate the following processes: filtering efficiency, visual search, multiple object tracking, complexity and resolution of the items, binding and grouping, WM resetting and updating.

Results that fall into the category of filtering efficiency had been previously examined. Vogel (2005) was able to demonstrate that low-capacity individuals find more difficult to filter out information that are not relevant and, subsequently, represented also the task irrelevant information. In fact, their CDA amplitude was similar in both filtering and non-filtering condition (i.e., when four relevant items were present). Moreover, individual filtering efficiency and VWM capacity correlated.

As for visual search, the role of VWM during this type of experimental procedure was clarified by a study from Emrich (2009) that measured CDA. In a typical visual scene paradigm, many items are present, and the participant has to identify targets among distractors. Moreover, since VWM has a limited capacity, the observer will not be aware of every stimulus displayed on the screen. In fact, the observer will only be aware of a small subset of objects at any given time (Hilimire et al., 2011). Notably, Hilimire and colleagues (2011) did a study to investigate both distractor and target processing during a visual search task, by analyzing several ERPs, such as N2pc, Ptc[1] and SPCN (i.e., the CDA). Results show that while the N2pc support the allocation of attention towards salient items and Ptc supports distractors suppression, the SPCN reflects continued processing of the target only, by representing the target itself in VWM (Hilimire et al., 2011).

In Emrich's study (2009) participants had to identify an upright "T" shape among other distractors, which were the same letter but rotated 90° and 180°. Before the search array, a cue appeared over the fixation cross to show to the participant in which direction he had to look. After completing the lateralized search task, participants had to complete a change detection task that they were not aware of. The most interesting result indicates that there was a sustained activity between 300-900 ms even during the visual search. The authors refer to this as "Contralateral Search Activity" (CSA), since the visual search task does not include a delay, and they conclude that VWM resources are employed while searching for a target (Emrich et al., 2009). Moreover, the amplitude of the CDA during the two tasks did not significantly differ, suggesting that a similar number of stimuli was encoded and held in WM in both tasks. However, as previously noted, the CSA had a later onset (i.e., from 300 to 900 msec) during visual search compared to the change detection task (i.e., 300 msec). This could potentially suggest that, during visual search, the encoding of items in VWM is more gradual. Thus, in compliance with Emrich (2009), the

---

[1] Hilimire et al., (2011) identified a positive ERP wave that was elicited contralaterally to the observed items. Since this ERP had a more temporal (e.g., more lateralized) distribution compared to the N2pc, it was named Ptc.

time that passed until the CSA reached its peak should be related to VWM capacity, and the CSA in high-capacity individuals should peak later. In fact, the correlation between VWM capacity (measured during change detection task) and the increase in the CSA amplitude was r = -0.84. Moreover, a significant correlation between search efficiency and VWM capacity was found (Emrich et al., 2009).

In accordance with Balaban and Luria (2019), VWM logically plays a major role in Multiple Object Tracking (MOT), in which subjects track a set quantity of identical moving stimuli that remain visible throughout the entire task. Drew and Vogel (2008) provided evidence that the CDA variability strongly predicts tracking performance. More specifically, the variations in amplitude that occurred between different numbers of to-be-remembered targets (i.e., rise from one to three targets) were the primary predictor of individual tracking ability. In compliance with the results that were found by Vogel (2005), this finding could suggest that individual variability in tracking performance could be due to filtering efficiency. More precisely, the observed differences could depend by how efficiently the visual system can separate target from distractors when tracking stimuli in a visual scene (Drew and Vogel, 2005).

Furthermore, Drew and colleagues (2013) set up a MOT paradigm and found that the CDA amplitude is also sensitive to dynamic changes in the number of targets during the task (Drew et al., 2008). So, it is possible to consider the CDA as online index of the quantity of stimuli that are being held in memory, even when the quantity varies across time. Moreover, when the number of distractors increased, the CDA amplitude remained unchanged but behavioral performance worsened, because the observer “swapped” target with distractors. By contrast, when the speed of moving objects was the parameter that increased, the CDA amplitude decreased because a target was being “dropped” (Drew et al., 2013).

The CDA may also be used to study how VWM capacity varies with changing object’s complexity and resolution. Behavioral evidence does not yield clear conclusions. In fact, some studies found complex objects to be more consuming of VWM capacity, while others found intact representations for complex items and stated that the observed decrement in performance was caused by processes that occurred after the retention interval (Luria et al., 2016). Luria and Vogel (2011) tried to solve this controversy using integrated objects to observe whether they would be represented as bound objects in

VWM or not, consuming additional storage. The experimental paradigm that the authors applied was the change detection task previously adapted by Luck and Vogel (1997). This study provided further supporting evidence that WM load is object-based and that WM maintains complex and integrated items (i.e. with a conjunction of characteristics) without costs. Thus, additional WM resources were required to maintain a polygon and a color as separate objects compared to when the same information was presented as combined objects, supporting the predictions made by the integrated object view (Luria and Vogel, 2011). The authors did a second experiment in which WM capacity consumption for color-color conjunction stimuli was measured. More specifically, participants were asked to decide regarding a change in the color of the polygon in polygons that had only one color, or in polygons that had a color-color conjunction (i.e., two different colors, so two decisions were required). Results indicate that accuracy was the same, however, CDA showed a slight conjunction cost by starting to decrease with the progression of the retention interval (Luria and Vogel, 2011). Notably, the integrated object viewpoint is also more coherent, according to discrete models of capacity (i.e., slot-based theories). Furthermore, a more recent study from Balaban and Luria (2015) showed that half of a polygon evoked the same CDA amplitude than the entire polygon, further supporting slot-based theories.

In accordance with the slots + averaging model (Zhang and Luck, 2008), one possibility is that VWM allocates more than one slot to items without a long-term memory trace. Therefore, polygons exhibit augmented CDA activity since they lack long-term memory traces, requiring multiple slots to be actively represented. Supporting this idea, it was discovered that the CDA for non-words is greater than for words, suggesting that complexity can stem from the absence of semantic meaning (Predovan et al., 2009). In conclusion, how VWM capacity is allocated remains an open debate in literature, although evidence seem to support the discrete slots theorization.

It is also possible to investigate the relationship between the CDA amplitude and the concept of "chunks". More precisely, as mentioned above, CDA amplitude appears to reflect the number of chunks (i.e., conjunctions of objects) rather than the number of different features. Therefore, CDA seems to be an adequate component to study online grouping in VWM paradigms (Balaban and Luria, 2019). Gestalt grouping principles have an important role in the interpretation of visual scenes and in perceptive processes.

Specifically, the brain pieces together the objects that are present in the world according to Gestalt principles such as, for example, proximity or common fate (i.e., Gestalt grouping cues), the latter states that when object move together, humans perceive them as a group. Luria and Vogel (2014) used a color change detection paradigm to study the relation between grouping and CDA. Specifically, in the critical condition a Gestalt grouping cue (i.e., common fate) was introduced, so the initial representation of the item would be updated and integrated into one unit. Thus, participants were asked to remember the colors and not the movement. On some trials the colors moved in different directions to give the idea that they were separated, and, on other trials, they moved in the same direction (i.e., common fate) or they moved independently and then "met" when landed on top of each other (i.e., proximity cue). Results indicate that common fate was the most effective grouping cue (Luria and Vogel, 2014), as demonstrated by the CDA amplitude that reflected that the items were grouped together. Proximity was less effective that common fate, however, high-capacity WM individuals were more able to group items under proximity condition. This study demonstrated how the integration was also based on movement history and not only on perceptual characteristics, since colors remained separate when they met but did not move together (Luria and Vogel, 2014).

Notably, binding does not seem to qualify as an obligatory process since VWM presents as quite flexible in terms of which objects characteristics should be held in memory.

Lastly, the CDA has also been proved to distinguish between updating and resetting, which are two different processes that can both alter VWM's representations (Balaban and Luria, 2019). When an item that is represented in WM changes, this can be handled in two ways, either by updating the existing representation to include the change (i.e., updating), or by discarding the old representation to create a new one that reflects the change (i.e., resetting). According to Balaban and Luria (2019 b), VWM can update a new representation only by resetting the old representation first. This theorization is corroborated by the results there were found by the authors in a study from 2019 (Balaban and Luria, 2019b). Specifically, a drop in CDA amplitude followed by a gradual recovery (i.e., resetting) is observable when a single shape moves coherently, and it is then divided in two separate parts. In a control experiment, the authors were able to demonstrate that separation is not even necessary nor sufficient to cause a sharp decrease in CDA amplitude. In fact, resetting can start also after object switching. Moreover, the drop in amplitude is only observable with resetting, and it is not present with updating processes (Balaban and Luria, 2019).

In conclusion, the CDA allows researchers to study numerous phenomena using quite a lot of different experimental paradigms. The CDA can give precise and informative insights about both VWM capacity and processes, and it qualifies as a very fundamental tool for future research, to better comprehend the role of working memory in various situations.

1.2.5 Precautions to adopt with experimental design and data analysis

When using the EEG to register the CDA, there is the need to control for eye movements and blinks. Even if the CDA is a parietal-occipital component that can be best observed at PO7 or PO8, or even at PO3 and PO4, blinks can indirectly effect EEG waveforms. Furthermore, blinking prevents the observer from seeing the visual scene on the display for about 300 ms. This issue seems mostly relevant for online processing paradigms which usually comprise long presentation time of the stimuli. Hence, online paradigms present more opportunities for artifacts (i.e., mostly blinks) to happen (Balaban and Luria, 2019). In addition, using Independent Component Analysis (ICA) during signal preprocessing does not completely solve the aforementioned issue.

Thus, it is appropriate to encourage the participant to avoid as more eye movements or blinks as possible during the instruction phase. It is also appropriate to insert a practice phase before the beginning of the experiment. By doing so, the observer will have an opportunity to get used to move his eyes as little as possible, with ongoing feedback from the experimenter. Alongside instructions, it is best to include a blank screen for 1-2 s after each response trial, to allow the participant to rest his eyes before continuing the task. Other practical tools include inserting breaks during the procedure. However, breaks must not be too long, otherwise participants will lose interest in the task and, most importantly, they will lose concentration. It is also wise to keep the observer engaged to the task by encouraging him during breaks. Thus, alpha waves that appear when the participant is tired will be less visible. Moreover, if the experimental design causes eyes to move frequently, an appropriate choice would be to restrict the movement direction. Lastly, examining EOG is highly suggested (Balaban and Luria, 2019). In conclusion, CDA does not reflect those artifacts, so changes in CDA amplitude are not caused by eye movements (Balaban and Luria, 2019).

Concerning the experimental design, if possible, it would be best to randomize locations of the stimuli across trials to register a clear CDA component, since using predetermined locations might attenuate the CDA (Balaban and Luria, 2019).

Furthermore, as previously mentioned, the strongest activity is in PO7/8 electrodes. Therefore, a common practice during data analysis is to average those electrodes with PO3/4 and P7/8, to make sure that no electrode pair is dominating the interpretation. However, this might attenuate the effects (Balaban and Luria, 2019). In addition, the traditional time window used starts at 300 ms after the memory array onset until the very end of the retention interval, but this may not be always applicable with online paradigms, which often require the time window to start only 200 ms after the main event (Vogel et al., 2005). Moreover, the trial should be divided in multiple time windows of interest.

To conclude, the CDA is a highly valuable marker that allows to study WM processes objectively and it will likely play a major role in the future, to better understand how VWM guides people’s behavior.

CHAPTER 2

2.1 ***CONSCIOUSNESS***

2.1.1 Definition of consciousness

The simplest way to define consciousness, according to the Webster's Third New International Dictionary (1966), is awareness of internal and external experience. However, the nature of consciousness has been discussed for centuries by scientists, theologians and philosophers. Notably, during late 19th century, the concept of consciousness was declared to be unscientific due to the rise of physicalistic reductionism[2]. Moreover, in the 1920s, Skinner theorized the radical behaviorism[3] and stated that every mentalistic concept had to be cancelled from psychology, therefore deleting about two-thirds of English words related to the subject. By the mid-1960s the behaviorism ceased to have such influence in psychological disciplines, especially in the United States. In 1989, Stuart Sutherland (1989) wrote that it would have been impossible to define consciousness, since it is such an elusive phenomenon. However, in the last 35 years, many advances were made while studying consciousness following several neuropsychological clinical cases and neurosurgeries (Seth, 2018). For example, split-brain studies (e.g., after callosectomy) revealed that each hemisphere is independently capable of perceiving items in the contralateral hemifield. Moreover, some systems that include the somatosensory and motor systems could be split in the same way, while others could not (e.g., emotion). Those findings challenged the idea that consciousness was unified in the brain, and this assumption was further discredited by the case of patient H.M., or Henry Moliason. Henry Molaison suffered from an acute form of epilepsy, so he underwent a bilateral removal of the medial temporal lobe in 1953. He then suffered from anterograde and retrograde amnesia and could not learn any new motor skill. However, the patient's working and semantic memory were not affected, and he could also acquire new unconscious (i.e., implicit) memories. These findings further challenged

---

[2] Physicalism assumes that everything is physical and refuses the mind-body concept. Therefore, concepts such as "consciousness" are not thought to exist (Taylor, 2010).

[3] Behaviorism is the study of laws related to observable behavior, without consideration for internal mental processes.

the concept of unified consciousness, since H.M. could acquire unconscious memories whilst his ability to form conscious memories was, by contrast, severely impaired (Seth, 2018).

Since the 1990s research about conscious processes has started to focus on the neural correlates of consciousness (NCCs)[4]. The search for NCCs was particularly accelerated by the improvement of neuroimaging and electrophysiological techniques such as MRI scanners and EEGs, that allowed to further locate processes or brain areas associated with specific conscious experiences or with being conscious at all (Seth, 2018). In fact, many experimental paradigms and sophisticated techniques were developed to study conscious experiences. For example, in binocular rivalry conscious percepts alternate whilst the sensory stimulus remains identical, allowing to compare conscious and unconscious conditions. Another example is given by masking paradigms that create the possibility of comparing subliminal and supraliminal presentations in many different sensory modalities. In the same period research also focused on studying the role of consciousness in different conscious states, both following brain damage (e.g., coma or vegetative state) and reversible (e.g., anesthesia). In this case, researchers tried to identify the brain regions that support being conscious and therefore should not be active when the person is not conscious.

Is it possible to state that defining consciousness is one of the most baffled problems in neuroscience (Chalmers, 1995). In fact, everyone intimately knows consciousness. However, there is not a single phenomenon that is harder to explain. In 1995, Chalmers (1995) introduced the idea that there is more than one problem related to consciousness. More specifically, there are several consciousness-related phenomena that have to be addressed so, according to the author, it is appropriate do distinguish between the “easy” and “hard” problems of consciousness. Easy problems concern with behaviors and functions typically associated with consciousness and they include allocating attention, controlling behavior, describing mental states, reacting to and discriminating stimuli, and volition. For example, according to Chalmers (1995), a person may say that a mental state is conscious when it is internally accessible or verbally reportable.

---

[4] Neural correlates of consciousness (NCCs) correspond to “the minimal set of neural events that is jointly sufficient for a conscious state” (Seth and Bayne, 2022).

Moreover, deciding to do an action is a conscious act when the action itself is deliberate. The hard problem of consciousness is, instead, experience. More precisely, the hard problem is constituted by the extreme difficulty of defining and understanding why and how a person can have subjective experiences, also labeled as “qualia”. In accordance with Chalmers (1995) experience is unique and differs from every person, and the explanation of this problem is the hardest since there is an explanatory gap, meaning no satisfactory explanation of experience in functional or physical terms exists (Seth and Bayne, 2022).

During more recent years, many theories of consciousness (ToCs) have been developed. Some theorizations directly focus on solving the hard problem, while other theories focus more on describing the easy problems, explaining conscious behaviors or functional properties associated with conscious processes.

2.1.2. Theories of consciousness

As mentioned above, when scientists began to study consciousness, researchers mostly focused on NCCs. Identifying NCCs has the advantage of being theory-neutral (Seth and Bayne, 2022), so researchers developed a common methodology and language to describe and study conscious phenomena. Due to the metaphysical neutrality (Klein, 2020), they were able to efficiently communicate, despite having different theoretical backgrounds among them. However, identifying NCCs posits several limitations. The most important limit seems to be differentiating the actual NCCs from the general neural activity that arises during conscious processing. More specifically, there are several papers in literature that highlight the presence of many obstacles that occur while trying to individuate NCCs in many experimental paradigms. As an example, De Graaf and colleagues (2012) emphasized that it is not possible to separate the neural prerequisites, the neural consequences and the neural substrates of conscious experience by using the neuroimaging techniques that are currently available. Despite these three elements are correlated to each other, they do not correspond to NCCs per se. It is indeed impossible, according to the authors, to extract elements solely related to consciousness (e.g., consciousness itself) from the subjective experience that the person is having.

Furthermore, Tsuchiya and colleagues (2015) highlighted that it is possible to both underestimate and overestimate the NCCs. More precisely, both report-based paradigm (i.e., paradigms in which the participant has to declare he was conscious about the stimuli) and no-report paradigms (i.e., absence of explicit report from participants) share the possible underestimation of NCCs. In particular, report-based paradigms frequently present experience without access, forgetfulness and non-reportability of the conscious stimulus, for example due to inattentional amnesia (Tsuchiya et al., 2015). These phenomena may cause researchers to underestimate NCCs, since some conscious experiences can be lost. Moreover, this can happen also in no-report paradigms because the stimulus may be experienced only if it has to be reported (Tsuchiya et al., 2015). As for the possible overestimation of the NCC, it is indeed a frequent phenomenon in report-based paradigms, because there can be an inclusion of post-perceptual processes and pre-perceptual processes, in accordance with De Graaf (2012). In addition, no-report paradigms are frequently subject to the inclusion of non-conscious processing, which can further cause NCCs to be overestimated. According to Tsuchiya and colleagues (2015), the best alternative to isolate NCCs is to utilize the binocular rivalry, as will be thoroughly described below.

Since defining NCCs entails those aforementioned limitations, the research focus has been steadily shifted towards developing an accurate ToC. The common aim of these theoretical approaches is to identify an explanatory link between the many aspects of consciousness and the neural processes associated with them (Seth and Bayne, 2022). Moreover, some theories assume that the explanatory gap, that constitutes the base of the hard problem, should be solved in order to clarify what consciousness actually is. By contrast, other theories do not specifically address this explanatory gap and focus instead on explaining the biophysical basis of consciousness (Seth and Bayne, 2022). It is also interesting to note that, as behavioral data are accumulating over time, many ToCs are emerging, instead of being ruled out. While this phenomenon can have positive consequences by gradually developing what will be an accurate and empirically proven theorization, it is indeed caused by the uncertainty that isolating conscious processes provokes. In addition, ToCs have different explanatory purposes, since researchers often do not agree on which phenomena a ToC should account for.

Before discussing the most relevant ToCs for the purpose of the study that will be presented below, it is fundamental to introduce some preliminary concepts related to consciousness. Firstly, there are two groups of states of consciousness: local states and global states. Local states (i.e., “states having qualia” or “conscious contents”) are defined by “what it is like” to be in them (Seth and Bayne, 2022) and include mood, autobiographical memories, emotions, volition and other factors. They even include the thoughts that arise when doing a particular action. Most importantly, these states are not independent from each other, but instead they are strictly connected to the scene in which the person can be found. By contrast, global states refer to the person’s overall profile and describe its behavioral responsiveness and level of arousal. More specifically, they include dreaming, wakefulness, coma, minimally conscious state and, as argued by some, psychedelic states that enables the person to reach higher levels of consciousness (e.g., with hallucinogenic substances). In fact, these global states are also identified as levels of consciousness, because the individual’s responsiveness and arousal significantly vary among them.

The second relevant distinction should be made between two different properties of consciousness: functional and phenomenal. The former mainly refers to the action that a certain situation or object consciously evokes. As an example, seeing a knife could enable different functions, such as cutting food or causing violent actions. The latter, instead, involves the actual experience (e.g., “what is it like?”). This distinction is quite fundamental because some ToCs focus only on one property, while some try to account for both of them. Notably, according to Seth and Bayne (2022), functional and phenomenal properties are most likely dependent from each other.

The third notable aspect is that different ToCs try to answer different questions regarding local states. As an example, there are local states, or “contents” that are always conscious, and some that cannot be conscious at all, and this can change according to the context. This fact can be empirically demonstrated with the usage of binocular rivalry, in which the conscious percept varies and alternates between the two images that are presented to the eyes.

Sharing Seth and Bayne’s (2022) perspective, the most relevant and influential ToCs can be organized in four groups: higher-order theories (HOTs), re-entry and predictive processing theories, Integrated Information Theory and Global Workspace theories. These theorizations will be described below, with a strong emphasis on Integrated Information Theory and, especially, Global Workspace theories.

In accordance with Brown (2019), HOTs try to describe the type of consciousness that refers to subjective experience (i.e., phenomenal consciousness). These philosophical theories assume that a first-order representation (such as an emotion, thought or perception) is not enough for conscious experience to arise (Brown et al., 2019). More specifically, to be conscious, a meta-representation of the first-order representation is needed. In fact, without this higher-order awareness, the first-order mental state would remain completely unconscious. However, Seth and Bayne (2022) highlight how HOTs are not entirely clear about how there are different types of phenomenal experiences. Moreover, the most common critique is that higher-order awareness seems too sophisticated to be actually plausible (Brown et al., 2019). Furthermore, the amount of events that are experienced every day is far too rich to be processed through high-order mechanisms. In addition, Lau and Rosenthal (2011) point out how these theories have been highly subject to criticism since they assume the involvement of the prefrontal cortex for the awareness of the perceptual information (i.e., higher-order representations). More specifically, this finding is criticized because prefrontal activity could solely reflect attention. However, the authors conclude that studies that found attention to be the reason of prefrontal activity do not draw the correct conclusions (Lau and Rosenthal, 2011), since attention does not account for variations in subjective reports of awareness during an experiment made with TMS (Rounis, 2010).

Re-entry and predictive processing theories consider top-down signaling to be a core and necessary factor in the generation of conscious experience. Whilst re-entry theories are proper ToCs that connect conscious perception to top-down processes (Lamme, 2006), predictive processing theories just provide explanations of brain and body functions, which can then be used to develop further accounts about the characteristics of consciousness (Howhy and Seth, 2020). Lamme (2000; 2006) studied visual processes extensively before elaborating the theory of Recurrent Processing (RP), to better establish how a person becomes aware of a stimulus after seeing it. More precisely, the author described a 3-step mechanism that include: feedforward sweep, localized recurrent processing and widespread processing. When the visual information is perceived by visual cortices in posterior occipital areas, the feedforward sweep is a process that rapidly enables the motor cortices to produce a potential response after they receive the input from the visual cortices. This process happens in just a few milliseconds and leads to the extraction of several features from the visual input such as, for example, color or motion. Notably, the feedforward sweep does not lead to a conscious experience. Then, the localized RP causes an exchange of information both within and between higher and lower areas. The last step, however, is the only one that allows the person to be conscious about the stimulus that he has seen. Specifically, the widespread recurrent processing allows the information to be shared with the areas that support executive functioning, such as frontal areas, through the frontoparietal network (Lamme, 2006). However, the author points out how he does not claim his approach to be ultimately correct. Instead, he suggests to strictly observe neuroscientific discoveries about consciousness, to be guided by them and to abandon traditional ideas about consciousness.

As for Tononi's Integrated Information Theory (IIT, 2004), the author developed a theoretical framework aimed at understanding the nature of consciousness. More specifically, Tononi (2004) aimed at exploring both the quality (or content) of consciousness, and its quantity (or level). According to IIT, "consciousness corresponds to the ability of a system to integrate information" (Tononi, 2004). Specifically, consciousness has two fundamental properties: integration and differentiation. The term "differentiation" refers to the fact that every system (e.g., the brain) has access to different conscious experiences, while the term "integration" denotes the degree to which a system's elements are unified and interconnected to produce a singular, cohesive experience. A system that comprehends a vast quantity of information is a system that has many different states and can distinguish among them effectively. Moreover, the different parts of the system are integrated and cannot be separated without losing their experience as a whole. Therefore, this concept emphasizes that the system is capable of generating conscious experiences as a whole, rather than as separate, independent components. For example, the experience of seeing a bird cannot be divided in several experiences of seeing wings, seeing eyes and seeing a beak. In addition, the degree of integrated information in a system can be quantified by Phi (Φ). When the value of Φ is high, the system is highly integrated and differentiated, and therefore it has a rich conscious experience. When the value Φ is very low or zero, the system is unconscious, or it lacks either differentiation or integration (Tononi, 2004). Moreover, Tononi defines the concept of "complex", which is a group of elements with $\Phi>0$ that is not part of another group of elements with higher Φ. According to this, Φ of a complex is used to measure consciousness. Moreover, IIT tries to solve the hard problem in a peculiar way, by deriving the requisites for the physical substrate of consciousness (PSC) from the indispensable characteristics of phenomenal experience (Tononi, 2016). More precisely, the theory identifies several axioms from which the characteristics of the PSC can be deduced. These axioms are intrinsic existence, composition, information, integration and exclusion, and from these axioms it is possible to define postulates about the PSC properties. The first axiom, which is intrinsic existence, theorizes that experience exists intrinsically. Therefore, the related postulate regarding PSC assumes that consciousness must be conceived in terms of "cause-effect power" (Tononi, 2016), meaning that consciousness is linked to the characteristic of the intrinsic cause-effect structure of a system. In other words, for something to exist, it must display a cause-effect power, meaning it must be able to change its state (Tononi, 2016). Secondly, the axiom of composition refers to the multicomponent aspect of experience. Thirdly, information comprises that experiences are composed of many aspects (e.g., qualia) that make every experience different from others. As for the integration axiom, as mentioned above, every experience cannot be divided in multiple parts because it is, by definition, integrated. Lastly, exclusion implies that every experience has its precise spatio-temporal boundaries (Tononi, 2016). IIT further describes the PSC, identifying a posterior cortical "hot zone"[5] as the cortical hub for consciousness to arise. The theoretical frameworks proposed by Tononi (2004; 2016) are several and the theory is not exempt from critiques. IIT implies that consciousness distributes on a spectrum and that it will be possible to establish the level of consciousness of a system (e.g., brain or a computer) by calculating Φ. However, calculating Φ in complex systems like the human brain is one of the major challenges of IIT, leading to a difficult applicability of the theory to research paradigms. According to Seth and Bayne (2022), IIT also poorly describes the relation between consciousness and other cognitive functions such as attention, memory or learning. Furthermore, IIT is criticized because it assumes posterior regions to be the hub for conscious processes, whilst there is evidence that found anterior cortical regions necessary for perceptual

---

[5] A broad range of cerebral areas that include occipital, temporal and parietal regions, and even the precuneus (Tononi, 2016).

consciousness (e.g., consciousness that arises from the subjective and unified experience of perceiving stimuli), as it has already been described above (e.g., Lamme, 2006).

Global Workspace Theories (GWT) are cognitive architectures that consider consciousness to have an explicit role (Baars, 2005). In fact, the first GWT (Baars, 1988) was theorized by Baars in 1998, and it focused on just the cognitive level. More precisely, the mental states that are conscious are the ones that are available globally to different cognitive functions such as memory and attention. To phrase it differently, GWT conceptualizes consciousness as a global workspace where information can be integrated, to then be broadcasted to the other brain systems. This enables efficient decision-making and problem-solving. Baars (2005) metaphorically equates GWT as a theater of mental functioning. Specifically, global workspace (e.g., consciousness) constitutes a bright light spot on a stage, and it is guided by attention, which is under executive control. The bright spot is, indeed, the only conscious part of the entire theatre, which remains unconscious. The main aim of consciousness is then to be distributed to an audience of networks which are present in the dark areas of the theatre. Therefore, the fundamental role of consciousness seems to integrate and coordinate the multitude of different networks that exist autonomously and independently in the brain (Baars, 2005). Without describing consciousness metaphorically, Baars theoretical framework can be summed as it follows. The majority of cognitive processes (e.g., attention) is unconscious, independent and operates in parallel. It is only when the information reaches the global workspace that it becomes conscious, due to it becoming integrated and shared across different brain systems because of higher-level processing.

Dehaene and Naccache (2001) elevate Baars theorization by formulating the Global Neuronal Workspace Theory (GNWT). Baars' GWT only described consciousness at a cognitive level, whilst GNWT dives into its neuronal aspect and characterization, adopting a more neuroscientific approach. Before presenting their theory, the authors highlighted three empirical findings that they evaluate as fundamental for consciousness. Firstly, cognitive processing can happen without consciousness. Evidence of this phenomenon derive from neuropsychological disorders that include prosopagnosia, neglect, agnosia, achromatopsia, split-brain patients and many other conditions (Dehaene and Naccache, 2001). An example is provided by Tranel and Damasio (1985), who made a study with patients suffering from prosopagnosia. Specifically, those patients exhibited a significant skin conductance response when looking at faces of person that they knew. By contrast, the same patients did not exhibit an electrodermal response when faces of strangers were displayed (Tranel and Damasio, 1985). The second empirical finding that was emphasized is that attention is a prerequisite[6] for consciousness (Dehaene and Naccache, 2001). The most relevant demonstration comes from patients with neglect, who fail to report items in the contralesional hemifield due to a difficulty in allocating attention towards it. Subsequently, they cannot be conscious of the contralateral items because they are unable to direct their attention towards them. Lastly, consciousness seems necessary for certain mental operations. More precisely, consciousness is required for information maintenance over time, to combine operations simultaneously as in the Stroop task, and for behaving intentionally (Dehaene and Naccache, 2001).

The GNWT postulate that the human brain's structure also includes a dispersed neural network, or neuronal workspace, with extensive interconnections that can link various specialized brain regions in a coordinated, yet flexible way (Dehaene et al., 1998). In accordance with Baars (1988), the global workspace provides a "communication protocol" that enables the modular cognitive systems to interact with each other, resulting in a conscious representation of the information. More precisely, there is a widespread and sudden activation across several brain areas. This phenomenon is defined "ignition" and leads to the generation of a conscious experience. The authors also claim that the process that allows the unconscious data to be made available to the global workspace can be defined as a top-down "attentional amplification" (Dehaene et al., 1998), therefore citing Posner's hypothesis of attentional amplification[7]. To enter consciousness, activity must be present for a sufficient period of time, to then become accessible to other processes. This "dynamic mobilization" is necessary; without it, the information will remain unconscious (Dehaene and Naccache, 2001).

---

[6] Note, however, that some stimuli can automatically and, most importantly, unconsciously capture attention (Yantis and Jonides, 1996). Consequently, attention can be oriented also due to unconscious bottom-up mechanisms, for evolutionary purposes.

[7] Attentional amplification is a phenomenon according to which allocating attentional resources causes more cerebral activation in some areas, therefore provoking an increase in their efficacy (Posner and Dehaene, 1994).

Furthermore, Dehaene and Naccache (2001) tried to describe the actual neural substrates of consciousness. Workspace neurons are present mainly in prefrontal cortices (PFCs) and anterior cingulate (AC). Specifically, the global neuronal workspace has the maximum level of connectivity due to the role of layers II/III of the cortex, which contain long and large pyramidal cells that can send long-distance axons, and these characteristics are mainly found in the PFCs (Dehaene et al., 1998). Notably, the neuronal workspace architecture is anatomically spread and is not only constituted by a single area or region. Otherwise, dynamic mobilization would not be possible. Dehaene and colleagues (2006) did several research on subliminal, preconscious and conscious processing, therefore proposing a taxonomy. Subliminal processing occurs when there is no reportability of the perceived information, and leads to very weak activation that dissipates in 1 to 2 msec. More precisely, this subliminal processing occurs when a limitation in bottom-up stimulus strength arises (e.g., the stimulus seems invisible). When preconscious processing occurs, the activation is stronger and spread to sensorimotor processors but cannot access the fronto-parietal circuit due to the momentaneous withdrawal of top-down attention. Lastly, during conscious processing, the fronto-parietal system is intensely activated, and the role of WM is mentioned (Dehaene et al., 2006). In addition, it is highlighted how there are both a late amplification of the gamma band and a synchronization in the beta range during conscious processing (Dehaene and Changeux, 2011) in experimental paradigms that measure the brain's activity through the EEG. To sum up the neural substrates, the role of the fronto-parietal network seems to be fundamental, and it is somewhat consistent with Lamme's proposal (Lamme, 2000; 2006) of a progressive accumulation of recurrent interactions. Furthermore, it is interesting to note that a tripartite distinction (i.e., subliminal, preconscious and conscious) has been suggested.

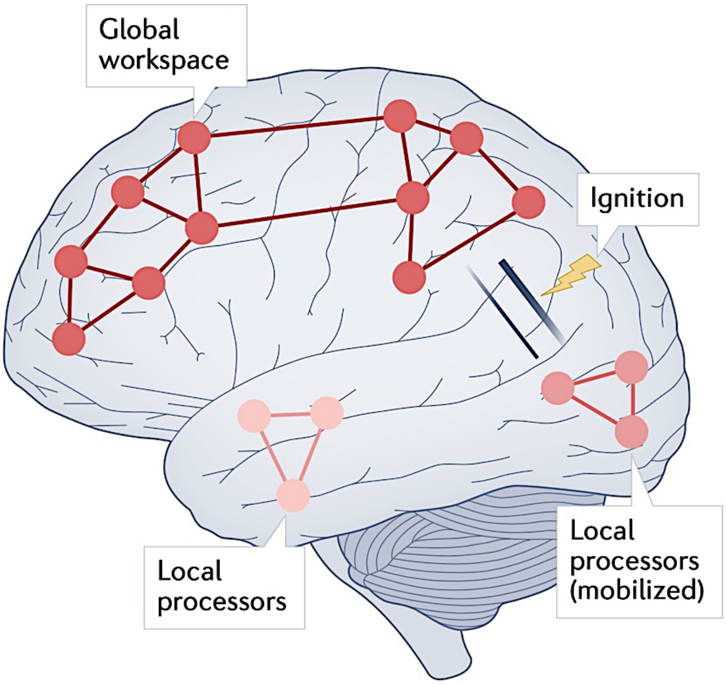


Figure 3. Visual schematic representation of the GNWT (Dehaene and Naccache, 2001), adapted from Seth and Bayne, 2022.

The theory has received some criticism since it does not give a satisfactory explanation for qualia (i.e., phenomenal consciousness). By contrast, GNWT seem to focus more on functional properties of consciousness, similarly to the majority of the ToCs. More specifically, the theory agrees on the multifaceted profile of conscious experience, but Dehaene and Naccache (2001) highlight that consistent neuroscientific research should be done before qualia can be fully understood and described. In addition, like HOTs, GTWs can be challenged by some evidence that claim that the involvement of anterior regions is caused by the behavioral response required in the experimental paradigm, and do not reflect conscious processes per se (Seth and Bayne, 2022). Nevertheless, GNWT might lead to clinical applications if conditions that display a reduced level of consciousness are displayed. Such conditions include schizophrenia, coma or simply anesthesia. Moreover, GNWT offers a precise and satisfying explanation of the relation between consciousness and the other cognitive functions (e.g., memory) and has had several neuroanatomical evidence that support its assumptions.

To conclude, ToCs are currently used, mostly, as narrative frameworks within the science of consciousness. While they guide the interpretation of neural and behavioral data, it is still uncommon for studies to be specifically designed with the goal of validating these theories. Although this approach is not problematic, future advancements will require experiments that directly test and clarify ToCs (Seth and Bayne, 2022).

2.1.3. The relationship between consciousness and visual working memory

In recent years, neuroscientists have been trying to clarify the relationship between WM and consciousness. However, this relationship has not been entirely clarified yet (Andrade, 2001). More precisely, it is unclear whether WM can retain unconscious information. Therefore, there is a possibility that also subliminal data could be manipulated by WM (Gambarota et al., 2022). To address the aforementioned issue, a metanalysis from Gambarota and colleagues (2022) will be thoroughly described.

Notably, GWTs directly consider the role of WM to be related to consciousness (Baars, 1997; Baars, 2005; Dehaene et al., 2006). Baars (1997) specifically writes “the contents of consciousness can be guided […] like a bright spot on the stage of working memory” and then states that WM is not entirely conscious. More precisely, the author assumes that only the events that are illuminated by the spotlight of attention become conscious. Notably, the spotlight is guided by the executive control, but the stage to which it is directed is the working memory itself, that has to maintain all the information (both unconscious and conscious) so that the information towards which the spotlight is directed can become conscious. Therefore, is it possible to conclude that, according to Baars (1997), WM has some aspects outside the focus of attention (Baars, 1997).

Over the years, many studies that tried to clarify the relationship between WM and consciousness were conducted. Moreover, the majority of the studies that were conducted adopted experimental paradigms that required the usage of VWM.

Some studies utilized the no-report paradigm to assess conscious experiences without making the participants report the stimulus that they were viewing. More precisely, Tsuchiya and colleagues (2015) hypothesized that the research for NCCs has been biased by reports, which cause the participants to engage in frontal activity to provide the answer. Therefore, the observed NCCs may have reflected the action of reporting an answer, and not the role of VWM during conscious processes (Tsuchiya et al., 2015). The authors then suggest using binocular rivalry to isolate the NCCs of stimulus processing, without including report as a confounding factor.

Moreover, to further confirm the theoretical framework that considers WM to retain also unconscious data (Baars, 1997), it is crucial to consider the findings of Soto and colleagues (2011). Their study directly examines whether WM can process unconscious information by utilizing briefly presented and masked Gabor patches. Following a delay period, participants were presented with another Gabor patch as a test. During the experiment, participants were tasked with discriminating the orientation of the target while maintaining the cue in memory, irrespective of its visibility, effectively performing a Change Detection Task. Finally, the observers had to rate from 1 to 4 how aware they were of the cue. The findings of this experiments demonstrate that it is possible for VWM to retain unconscious stimuli. Moreover, authors proved that results were not caused by unconscious priming mechanisms (Soto et al., 2011). The fact that WM manipulates information that is not directly present in the focus of attention is also consistent with state-based models of WM (LaRocque et al., 2014), which assume that the information represented in STM can exists in multiple different states, and that there is a common store for STM and LTM, although it has different states of activation (Cowan et al., 1995). These models emphasize the role of the focus of attention, highlighting how conscious awareness of the stimuli occurs when attention is directed to the stimuli. However, LaRoque and colleagues (2014) claim that the focus of attention may not be necessary for consciousness to arise, which is exactly the case in the study from Soto et al., (2011). Nonetheless, these state-based models quite differ from the memory models proposed by Baddeley (1974; 2000) or by Atkinson & Shiffrin (1968), despite the fact that they both assume WM to have a limited capacity, in accordance with slot-based theories (LaRoque et al., 2014).

Furthermore, the findings relative to WM being able to manipulate unconscious information have been severely criticized, at both methodological and theoretical levels. Notably, Stein and others (2016) pointed out that, in the study from Soto (2011), the metric used is not adequate to measure bias-free detection sensitivity. More precisely, the subjective ratings of the cue awareness that were adopted by Soto et al., (2011) can be prone to response bias. This means that the participants could not give an objective report of their stimuli awareness, therefore deciding to claim that a stimulus was invisible, when it was actually only partially visible. According to Schmidt (2015), subjective visibility judgements do not constitute a reliable methodology to discriminate between conscious or unconscious processing of the stimuli. Thus, evidence that claim that WM can be unconscious are lacking and inconclusive (Stein et al., 2016).

It is possible to conclude that, to this day, whether WM is unconscious is still controversial and undetermined. Therefore, the metanalysis written by Gambarota and colleagues (2022) aims at establishing the most relevant moderators in previous studies that tried to address that relationship, to then provide future directions for researchers that will try to clarify this apparently hard-to-solve issue. In particular, the authors tried to understand why the heterogeneity between previous studies is high, so the methodological heterogeneity in the included studies is analyzed (Gambarota et al., 2022). Some studies eligibility criteria included measuring VWM specifically, the behavioral performance had to be assessed through direct measures (e.g., Change Detection Task) and the to-be-remembered item had to be manipulated, at the time of encoding, in order for the observer not to be aware of it. 13 papers and 2 PhD thesis were analyzed, leading to a total of 38 different effect sizes (Gambarota et al., 2011). The authors chose to calculate a meta-regression to explain the estimated heterogeneity. More specifically, the moderators were the blinding technique[8] used, the type of WM task, the duration of the target presentation and the average number of trials for each participant. As for the blinding technique, there are several types of it that were adapted in the included studies: Continuous Flash Suppression (CFS), Attentional Blink (AB), Backward Masking (BM), Sandwich Masking (SM) and Metacontrast Masking (MM). CFS, firstly introduced by Tsuchiya and Koch (2005), involves presenting a rapid-changing high-contrast image pattern to one eye, while a static or slow-changing image is presented to the other eye. The image that changes frequently suppresses awareness of the target image. This technique has the advantage of reducing negative afterimages (Tsuchiya and Koch, 2005). Differently, AB is a phenomenon that was identified by Raymond and colleagues in 1992, and it corresponds to the inability of detecting a second target if it appears within 200-500 msec after the first target (Raymond et al., 1992). This occurs because the brain is still processing the first target, so there is a small attention lapse that causes the second target

---

[8] Blinding techniques are used in research to suppress the observer's awareness of the stimuli that are being presented during the experimental task (Kim and Blake, 2005; Breitmeyer, 2015).

to be easily missed. BM is a phenomenon where the perception of the target is either suppressed or hindered if the target is followed by another stimulus (i.e., “mask”) that typically appears after a short time since the presentation of the target (Breitmeyer and Ogmen, 2000). In BM, the mask spatially overlaps the target. MM is, instead, a type of BM in which the mask does not overlap the target spatially, but it is usually surrounds the target, creating a perceptual effect where the target seems to fade (Breitmeyer and Ogmen, 2000). Finally, SM involves the presentation of two masks, one before (i.e., “forward mask”) the target stimulus and one after (i.e., “backward mask”) the target stimulus (Nakano and Ishihara, 2020). As previously mentioned, another fundamental aspect is the type of experimental paradigm used to measure VWM. VWM paradigms used are the Change Detection Task, Delayed Detection (DD), Delayed Match-to-sample Task (DMS) and Delayed Estimation Task (DET). Change Detection Task has already been described above and it was adapted by Dutta et al., (2014), King et al., (2016), Pan et al., (2014), Soto et al., (2011) and Tagliatela Scafati (2019rep). In DD participants must detect a specific stimulus after a delay period and not immediately, so researchers can analyze the delay period to gather information about how data are maintained in WM over short periods of time. This paradigm was used by Bergström and Eriksson (2014, 2015, 2018), Taglialatela Scafati (2019a, 2019b, 2019c). The DMS, utilized by Bergström and Eriksson (2015, 2018) and Nakano and Ishihara (2020), begins with a presentation of the sample stimulus to the participant. This sample stimulus is then removed during the delay period and, in the matching phase, the observer has to declare which of the newly introduced comparison stimuli match the original sample stimulus. Finally, DET has the same structure of the other WM tasks (e.g., target presentation, delay interval and WM task). However, in the WM task, the participant is asked to estimate or reproduce the specific feature of the stimulus, to assess the precision of memory representations. The DET was adapted by Trübutschek et al., (2017, 2019a, 2019b).

The results of the metanalysis confirms the presence of an unconscious behavioral WM effect (Gambarota et al., 2022). The majority of the estimated heterogeneity, according to the analysis, can be attributed to the author’s model, since the authors did choose very diverse experimental approaches. Moreover, the blinding paradigm and the WM paradigm seem to have had a great influence on overall heterogeneity. More precisely, CFS appears to be the most effective blinding technique among the others. CFS is also the technique that leads to the lowest percentage of discarded trials, since it has a strong effect that effectively prevents the participant from consciously seeing the target (Gambarota et al., 2022). Furthermore, the DET (Trübutschek et al., 2017, 2019a, 2019b) has the highest effect-size, which appears curious since it probably is the most complex tasks among the tasks that were included. However, the stimulus complexity was very low in DET (e.g., a white square) in the studies from Trübutschek et al., (2017, 2019a, 2019b). Notably, according to Song and Yao (2016), easier stimuli like Gabor patches are unconsciously processed but with higher accuracy compared to more complex stimuli (e.g., faces). Hence, in future studies, DET should be performed with more complex stimuli to understand its real impact on VWM (Gambarota et al., 2022). According to the authors, it is also important to consider how much the observers are actually involved in the WM task, so it seems necessary to give appropriate instructions before beginning with the procedure. More specifically, it would be appropriate to actively engage the observer to the task when giving instructions, to increase accuracy by facilitating the engagement of attention towards the task. In addition, there is one last factor that could not be included in the meta-regression: consciousness operalization. The majority of the analyzed studies adopted a subjective approach by specifically asking the participant to declare his level of consciousness through specific assessments, usually presented between the retention and test phases. To obtain more robust results, consciousness levels should be measured objectively, to also avoid overestimation of unconscious processing (Gambarota et al., 2022). Another criticism was raised by Stein (2016), who claims that the unconscious WM effect can be attributed to the participant making a guess when presented with an unconscious target and then retaining that unconscious guess, without requiring any actual unconscious process (Stein et al., 2016). Therefore, it is necessary to get precise evidence that the unconscious and unseen data are genuinely retained (Gambarota et al., 2022). This issue (i.e., "conscious guess maintenance") was directly addressed by Trübutschek et al., (2017, 2019b) by applying multivariate pattern analysis (MVPA) in the delay period, and the authors did not find any proof of the pattern.

To sum up, Gambarota and colleagues (2022) suggest a more homogenous approach for future studies, to compare results more easily and, probably, to find more consistent results about unconscious processing during WM processes, which already appear to be existing.

2.2 ***BINOCULAR RIVALRY***

2.2.1 Description of binocular rivalry

Usually, the brain integrates the different retinal image that come from each eye to create a coherent representation of the visual scene. This process is called binocular superimposition, and it is crucial for depth perception and to create a plausible visual experience (Carmel et al., 2010). However, when the two images are too dissimilar, the brain alternates every few seconds between perceiving one image, and then the other, rather than combining them into a single coherent representation. This phenomenon is known as binocular rivalry and occurs because the brain is presented with two images that excessively differ from each other and, therefore, is unable to combine them. Subsequently, a perceptual competition (*Figure 4*), where is it only possible to see one image at a time, manifests (Carmel et al., 2010).

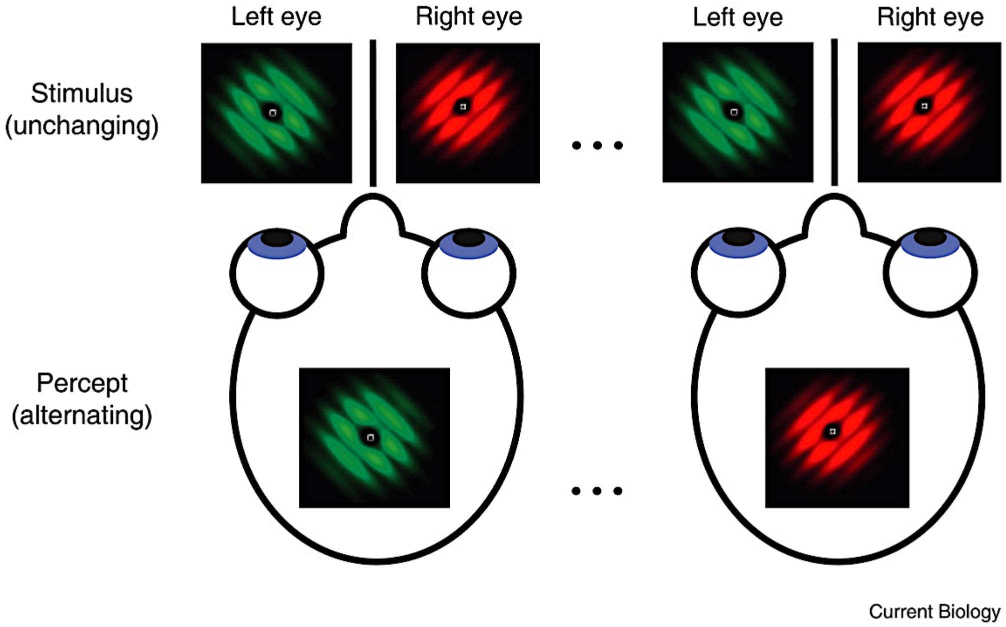


Figure 4. Visual adaptation of the binocular rivalry phenomenon. This image, adapted from Clifford (2009), shows how the two different percepts alternate, rather than being combined to create a single coherent representation of the visual scene.

Binocular rivalry can be provoked by a wide range of stimuli. For example, some dimensions of the stimuli can cause rivalry to occur, such as contrast, luminance, color, size or form, and others (Blake, 2001). Furthermore, rivalry can be evoked by both simple and complex stimulus differences. However, the stronger the competition is (e.g., the more different the images are), the faster the rivalry alternations are. Notably, not every difference between two images can cause rivalry to occur.

There are some spatial and temporal properties of binocular rivalry that are worth mentioning. As for spatial properties, when the two stimuli are too large, the observer frequently experience the so-called "patchwork" rivalry, otherwise called "piecemeal" rivalry (Blake, 2001). This phenomenon causes the perception of periods of mixed dominance, with portion of the view coming from both eyes. Another spatial property concerns perceived size of the images. More specifically, the spatial extent of rivalry is determined by retinal image size and not perceived size, because of Emmert's Law, that assumes the perceived dimension of an object to increase with perceived distance, when the retinal image is fixed (Blake, 2001). Moreover, Blake (2001) even describes how the transition from the dominant stimulus to the other happens. Precisely, the switch does not abruptly happen. In fact, it seems that the non-dominant stimulus sweeps the other from conscious awareness in the form of "spatially coherent waves of dominance" (Blake, 2001). In addition, examining temporal properties, the two rival images must be presented simultaneously for several hundred milliseconds (e.g., at least 200 msec) for binocular rivalry to occur (Blake, 2001). Moreover, it is impossible to predict the duration of individual periods of dominance and suppression of both images, and the rate of alternations increases with complexity and strength (e.g., contrast) of the rival stimulus. Predominance is one of the most important indices in binocular rivalry processes, and it measures how often one image is perceived, compared to the other. In addition, predominance it is usually expressed in ratio or percentage. Furthermore, shifts in dominance between the two stimuli are partly driven by neural adaptation, according to which the signal from the currently dominant stimulus gradually weakens over time, allowing the previously suppressed stimulus to take over and becoming the dominant one (Blake, 2001). Finally, it is important to note that there can be top-down influences on rivalry. As an example, attention can bias perception, favoring one of the two images. Nevertheless, motivational and cognitive factors could also bias the observer's perception.

In 2006, Tong et al., (2006) investigated the neural bases of binocular rivalry by analyzing several fMRI and EEG studies about the phenomenon. More precisely, the authors found that the ventral temporal cortex reflects the participants awareness in rivalry processes. Moreover, strong rivalry modulations have been found in the primary visual cortex (V1) and in the lateral geniculate nucleus (LGN). In addition, researchers have found a correlation between perceptual switches and right lateralized frontoparietal areas (Sterzer et al., 2009).

2.2.1 Usage of binocular rivalry techniques to study unconscious processes

In accordance with Tsuchiya et al., (2015), binocular rivalry paradigms provide an excellent tool to examine unconscious processing and, especially, to isolate NCCs and therefore excluding neural correlates of basic stimulus processing (Tsuchiya et al., 2015).

Nevertheless, Carmel et al. (2010) provided some guidance about how to successfully elaborate a binocular rivalry display. The most important aspect is to maintain stable vergence to always have a fixed image that is falling on each fovea. Therefore, if the images presented to the two eyes differ completely, vergence will be disrupted and, therefore, it will be impossible to evoke rivalry processes. Subsequently, the display should present some identical elements for the two eyes, in addition to the two different images, to maintain a stable gaze (Carmel et al., 2010). The most common way to stabilize vergence is to include both a frame around the images and a fixation cross, or dot, in the centre of the image. To reduce horizontal eye movements, it is also possible to include a textured bar on each image side. Finally, when the methodological aspects of the study prevent the researcher from using bars on the side of the images, it is possible to include an image that is further from the stimulus, or nonius[9] lines (Carmel et al., 2010).

The three most popular, straightforward and inexpensive methods for creating a binocular rivalry display are using color goggles, using a mirror stereoscope or using prism goggles.

Firstly, before adopting one of these techniques, it is necessary to prepare the stimuli for the upcoming experiment. The easiest and most practical way is to present images side by side on a monitor and change their characteristics in accordance with the technique that will be used, as will be described below. According to Carmel (2010), researchers prefer using red-blue (or red-green) cellophane goggles, since they are

---

[9] Nonius lines are lines that center the image from every direction (Carmel et al., 2010).

inexpensive and easy to find at most stores. When utilizing this type of goggles, it is necessary to adapt the images to the monitor and to change their color setting. More specifically, one image has to be displayed by the monitor's blue (or green) gun, while the other has to be displayed by the red gun, since every image will be passed by only one lens. The images should have the same information (e.g., fixation dot and frame) so stable vergence can occur. Moreover, these identical characteristics should be presented in a color that both lenses can let through, which is black. Using red-blue goggles has some advantages, including being easy to adjust and inexpensive. Moreover, they are compatible with MRI recordings and do not require neither the head to be stabilized, nor adjusting the view for each participant. By contrast, they are not exempt from disadvantages. As an example, chromatic images cannot be used, and they are not compatible with most current mobile eye-trackers. In addition, lenses are not perfect, and each eye will see also the image presented to the other eye. Consequently, it will be impossible to claim that the image that was suppressed was completely unseen (Carmel et al., 2010).

As for mirror stereoscopes, they comprise four different mirrors (*see figure 5A*). It is also possible to create one, rather than buying it pre-made. To do so, it is sufficient to arrange two mirrors close to each eye, angling each at 45° relative to the line of sight (use a chinrest to prevent the participant's head from moving). Then, two additional mirrors need to be placed on the outer sides of each of the initial mirrors, angled at 45° to face the stimuli (*see figure 5B*). Mirror stereoscopes differ from red-blue goggles since they require to be adjusted with each participant to obtain stable vergence, since everyone's eyes are different. In addition, it is fundamental to ensure that every eye is seeing just one image, and not a piece of the other image too. This is made possible by placing an object that divide the two hemifields, such as a piece of cardboard, in the middle of the mirror stereoscope (Carmel et al., 2010). Notably, however, the participant must not see the carboard when looking through the mirror, so it is best to avoid a cardboard made with a shiny material that will reflect light, and to prefer a matte color. Moreover, it is recommended to prepare one image to be used for calibration purposes, before starting the experimental procedure. Using a mirror stereoscope has several advantages: chromatic images can be utilized, the stimulus preparation is quite simple and straightforward, and, most importantly, the images cannot "bleed" into each other since they are completely separated, unlike with the previously described goggles. In addition, mirror stereoscopes can be used with eye-tracking, EEG, magnetoencephalography (MEG), near infrared spectroscopy (NIRS) and non-invasive brain stimulation (NIBS) techniques. Mirror stereoscopes, however, present some negative aspects too. As it has already been mentioned, they need to be adjusted for each participant, resulting in a time loss during the experimental paradigm, which can tire the participant even more and this could lead to less precise reports due to a disrupted attentional engagement of the observer. Moreover, the presented stimuli need to be small since there only is half of the monitor that can be used for each image. Finally, mirror stereoscopes are not compatible with MRI recordings. However, nowadays, this problem seems less of a burden since precise source-localization techniques have been developed and can be used with EEG recordings, without the need of using MRI.

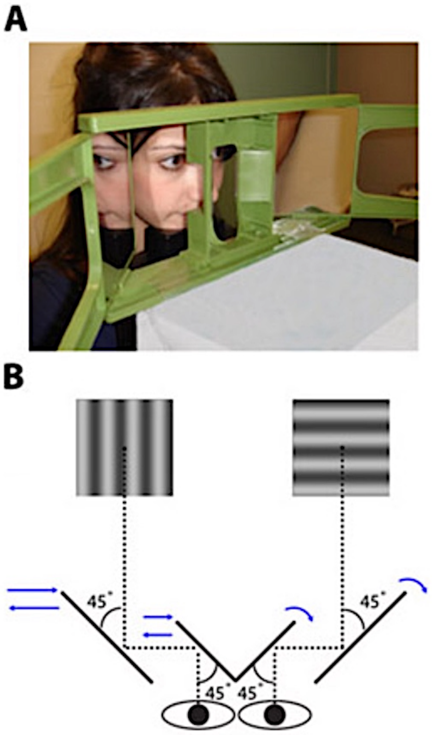


Figure 5. Illustration of a mirror stereoscope, adapted by Carmel et al., (2010). Figure A gives an example of the technique with its four mirror is given. Figure B represents the technical details of the mirror stereoscope. Black lines represent the position of the four mirrors, while the dotted lines represent line of vision. Possible adjustments that can be made are described by straight and curved arrows.

Lastly, prism goggles are a combination of the two aforementioned techniques. More precisely, they are goggles with prismatic lenses. Therefore, images are presented side by side as with a mirror stereoscope. Each prism bends light, causing objects that are actually to the side to appear directly in front. When two prisms are oriented in opposite directions, they create the illusion that two images, which are physically side by side, overlap in space, similarly to a mirror stereoscope. Notably, a divider is still needed, but the distance and size of the display do not have to be adjusted each time (Carmel et al., 2010). Using prism goggles has all the advantages of a mirror stereoscope, but prism goggles can also be used in MRI settings. As the other techniques, prism goggles are not exempt from disadvantages such as requiring head stabilization and causing distorted images when large stimuli are presented. Therefore, as with the mirror stereoscope, it is recommended to use small stimuli.

To conclude, binocular rivalry is an efficient technique that must be tailored to the researcher's needs to ensure that the measurement of the variables of interest occurs. Interestingly, it is an excellent paradigm to study unconscious processes, as mentioned above, and also by Tsuchiya (2015). More precisely, it allows researchers to measure the unconscious processing of one of the two images. For example, when image A and image B are presented, the subject may declare to see image A. Hence, it is possible to investigate the neural processes associated with the condition in which the observer declares that they are not seeing image B, despite it being presented to their eye.

CHAPTER 3

3.1 ***Goals and hypothesis***

This chapter is dedicated to outlining the purpose of the research and, therefore, a detailed representation of the methodology that has been employed will be given.

The experiment was conducted in the electroencephalography laboratory at the Department of Developmental Psychology and Socialisation (DPSS) at the University of Padova. Notably, this research has been conducted in collaboration with Professor Roy Luria from Tel Aviv University. The present study aimed to investigate whether visual working memory (VWM) is capable of retaining unconscious information. To explore this, participants underwent EEG recordings while engaging in a binocular rivalry paradigm, during which they discriminated between different stimuli. In addition to the EEG data, behavioral responses were collected, with participants required to indicate which type of image they perceived by pressing one of two designated keys on the keyboard. The technical aspects of the methodology employed in this study will be discussed in detail in the following sections.

It is important to provide an overview of the key concepts related to VWM, CDA and binocular rivalry, along with a detailed description of the main hypothesis of the study. As mentioned above, the goal of this experiment was to provide clear-cut evidence to the ongoing debate regarding the possibility that VWM can hold also unconscious representations. The CDA was chosen as the primary measure, as it is well-suited to assess whether items (both conscious and unconscious) are being retained in VWM. According to Luria and colleagues (2016), the CDA reliably reflects the number of items that are being held in VWM, since its amplitude rises as the number increases, up to an asymptote. Thus, since the CDA serves as a valid indicator of VWM capacity, it can be inferred that it is possible to deduce whether a person is maintaining an item in WM by looking at the CDA amplitude. More precisely, if a participant is consciously aware of an item on the screen, the CDA will be elicited, making it a valuable tool for studying unconscious stimulus processing. According to our hypothesis, the CDA should be elicited even when the participant is unaware of the stimulus, which is the case with binocular rivalry. Specifically, in a binocular rivalry paradigm, two items are presented to the observer's retinas, but the conscious percept alternates between the two, therefore causing one item to have a conscious representation in WM, while the other remains unconscious (Tsuchiya et al., 2015).

In the experiment, two stimuli were presented on a monitor, positioned to the right and left of a central fixation cross, while participants viewed the stimuli through a stereoscopic mirror to induce binocular rivalry. During this period, EEG signals were recorded. Based on the assumption that each stimulus appearing on one hemifield is processed by the contralateral visual areas, an activity that reflects stimulus processing is expected for both cerebral hemispheres in trials were the two stimuli appear at the same time. Specifically, the stimulus presented to the right of the fixation cross projects to the left hemisphere, and vice versa. Consequently, if VWM can maintain both stimuli (i.e., the unconscious and the consciously reported stimulus), the amplitude of the two ERP waves will be identical. Thus, the resulting CDA, which is derived from subtracting the ipsilateral wave from the contralateral wave (Luria et al., 2016), will equal zero. If the resulting CDA has an amplitude of zero, this would serve as evidence supporting the unconscious processing of information by the VWM. More precisely, the binocular rivalry condition ensures, as previously mentioned, that the subject is aware of only one image from the two images presented simultaneously on the retinas. Having participants indicate, via keyboard, which stimulus they are consciously perceiving allows us to assess whether unconscious processing of the other stimulus is occurring. If the CDA amplitude is zero at that moment, it would suggest that VWM is maintaining both the consciously perceived and the unconscious stimulus, providing evidence for unconscious processing in VWM.

To sum up our core hypothesis (H0), if the CDA amplitude equals 0, this can be considered evidence that the participant's VWM is retaining both the conscious and the unconscious stimulus, thereby indicating that VWM can maintain unconscious information.

3.2 ***Methods***

3.2.1. Participants

18 participants were recruited by contacting them through e-mail. Specifically, participants were either psychology students that voluntarily agreed to partake in the

experiment, or personal acquaintances. The only exclusion criteria included: 1) suffering from a psychiatric or neurological condition; 2) suffering from a cardiovascular disease; 3) taking medicines that have effects on the central and/or peripheral nervous system.

Specific attention was devoted to people suffering from frequent migraines, especially migraines with aura. It is well known, in both medicine and neuroscience, that stimuli with a high contrast, such as the ones that were used in this experiment, can trigger a migraine attack (Shepherd, 2000). Moreover, striped visual patterns have the same effect (Marcus and Soso, 1989). Therefore, suffering from migraine with aura or having frequent migraine attacks constituted the most important exclusion criteria. We also devoted the same attention to epilepsy, to ensure that participants did not suffer from it.

Every participant gave his informed consent according to the Helsinki declaration. Furthermore, the experimental procedure was approved by the local ethical committee for research and, therefore, the research was conducted according to the provided guidelines. Anonymity was guaranteed by assigning a numerical code to every participant. Lastly, every participant could interrupt the experiment at any giving moment due to any possible reason (e.g., discomfort, for example).

Finally, there was not any type of reward (e.g., monetary compensation) for participating in the experiment.

3.2.2 Materials

The stimuli that had to be discriminated during the experimental trials were Gabor patches (*Figure 6*). Gabor patches are sinusoidal gratings which, in this case, are located into a circle. More precisely, these Gabor patches comprise vertical or horizontal lines with a high contrast, so these types of stimuli should alternate frequently (Blake, 2001). Furthermore, the stimuli were black and white to maximize contrast, so a stronger binocular rivalry phenomenon could occur.

In details, the measurement of the stimuli did not vary across trials and was height x length. Those stimuli were presented on a grey screen either in pairs, one on each side of the fixation cross, or individually. The reason why a fixation cross was placed on the monitor is to stabilize vergence, as suggested by Carmel et al., (2010).

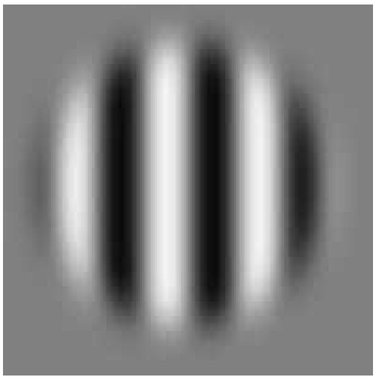


Figure 6. Example of a black and white Gabor patch, on a grey screen.

3.2.3 Experimental procedure

Upon arrival at the electroencephalography laboratory at DPSS, participants were provided with an informed consent to ensure they fully understood the research methodology and data collection methods. Special emphasis was given on explaining that all the collected data would remain anonymous, and that the participant could withdraw from the procedure at any given moment, even without giving any specific reason to do so. Afterwards, head circumference of the participant was measured with a tape measure to decide which EEG cap size was best for the participant's head. Usually, it is best to avoid a cap that is too small, otherwise the observer could suffer from migraine during the experimental procedure or afterwards. The EEG system was a 64-channel system with active electrodes, which required an adequate preparation of the participant's scalp and face. Once the appropriate EEG cap size was selected, all the 64 electrodes were placed on it. Moreover, in order to make the subjects feel safer and more at ease with the equipment used, the syringes were prepared in front of them, so it was possible to establish that the syringes were clean and that they had a blunt needle, which could not hurt or harm in any possible way. Then, the procedures involved in preparing for the experiment were explained, and the participant was seated in the experimental cabin. The participant was also asked to remove earrings and not to wear contact lenses, since they make eyes blink a lot more than regular glasses, introducing significant noise on the EEG signal. However, the observers could wear glasses after the cap montage was done. This is particularly relevant for our experimental purposes, since Gabor patches can particularly cause eye strain when staring at them for prolonged periods of time. In addition, clear vision was desirable to avoid any inaccuracy while detecting the stimulus orientation.

As for the scalp preparation, the EEG cap was placed and then centered on the head by calculating the distance from the two preauricular points and from nasion to inion. Consequently, the scalp was adequately prepared by lightly scrubbing it with the blunt needle, and by placing the Easycap SuperVisc gel under the electrodes, to lower impedance below 10 kOhm, if possible. Furthermore, the extended 10-20 system was used, with the reference electrode placed on the left earlobe. In addition, electrode TP9 and FT10 were placed on the side of, respectively, the left eye and the right eye to record saccades (e.g., horizontal eye movements). Moreover, an electrode typically placed on TP10 was placed on the right earlobe, and another electrode usually on FT9 was placed above the left eye to detect blinks. Since the CDA is a parieto-occipital component that is best observable at PO7 and PO8, special attention was directed towards the scalp preparation of these specific areas, to maximize lowering impedance. Moreover, the participant was always instructed to limit blinks, as they prevent the observer for seeing the visual scene for about 300 ms, as noted by Balaban and Luria (2019). Notably, however, the type of stimulus used can cause frequent blinking due to its high contrast. Hence, avoiding blinks for more than a few seconds was reportedly too difficult for the participants.

After the preparation, the experiment began and it was divided in three phases: calibration, practice, and the experimental trials. Before describing each stage, it is appropriate to thoroughly outline the experimental setup. As depicted in Figure 2 and Figure 3, a chinrest was placed in front of the monitor to ensure the participant's head was stable across the experimental trials. In fact, the observer had to put their chin on the chin guard to see through the stereoscopic mirror. Moreover, a piece of carboard was attached to the mirror, in order to ensure that each eye was looking at just one stimulus in one single visual hemifield, following the methodology described by Carmel et al., (2010). In addition, the piece of carboard was beige in color, to minimize light reflection.

Figure 7. Frontal view of the experimental setting. The chin guard helps to support the participant's head for the duration of the procedure. It is also possible to see the stereoscopic mirror attached to the chinrest.

Figure 8. This image shows a side view of the stereoscopic mirror to see how the piece of cardboard was attached to the rest of the equipment.

The calibration phase varied across participants, as each had to state whether they were viewing just one circle with horizontal or vertical bars. More precisely, two Gabor patches were presented on the monitor, one on each side of the fixation cross. When looking through the mirror stereoscope, the two Gabor patches had to perfectly overlap to allow binocular rivalry to happen. During this phase, the experimenters adjusted the position of the two stimuli according to what the participant was describing. For instance, many observers reported to see stripes or black dots on one edge of the screen when looking through the mirror and, therefore, the stimuli's position had to be tailored to meet the participants needs.

Once calibration was complete, the practice phase began. The participant was asked to say out loud if the viewed bars were horizontal or vertical, while pressing the corresponding button on the keyboard. More precisely, the key to be pressed were either "M" or "Z". Moreover, if the participant was assigned an odd numerical code (e.g., participant 011), the letter "M" had to pressed when seeing horizontal bars, while the letter "Z" when seeing vertical bars. Vice versa, if the participant was assigned to an even numerical code, the letter "Z" had to pressed with horizontal bars, and the "M" with vertical stripes. These two conditions were created to ensure between-participants randomization.

Lastly, the actual experimental procedure started after the practice phase. The design consisted of 4 blocks of trials of the duration of 8 minutes each, for a total of 32 minutes. The task remained the same of the practice phase, except the participant was no longer required to verbalize the responses given. Participants were required to press either the "M" or "Z" key to indicate whether they perceived horizontal or vertical bars while viewing the stimuli through the stereoscopic mirror. Additionally, each trial presented different combinations of Gabor patches, with variations in the width of the stripes. For instance, some patches featured large, widely spaced stripes, while others contained stimuli with very narrow bars. The precise duration of the experiment differed among participants since 3 breaks were included. After finishing each block, the experiment paused, allowing the participant to take a break as long as needed to rest the eyes. As a matter of fact, staring at frequently changing Gabor patches was reportedly challenging and tiring for some participants. However, this feeling is entirely subjective, and other participants did not even need to take a break between the blocks. Nevertheless, we decided to encourage the observer during breaks in order to keep their attention engaged to the task and to reduce alpha waves, which are caused by tiredness, during EEG recording (Balaban and Luria, 2019). Additionally, we did not allow breaks to be too long since there would have been the risk of losing focus on the task.

Finally, certain trials displayed only one stimulus, either on the right or on the left of the fixation cross, to create a control condition. By doing so, it is possible to observe whether we can extract a clear CDA signal that differ from the one that arises when two stimuli are presented simultaneously. The latter condition is, according to our hypothesis, the condition in which the CDA should equal 0. However, when looking at just one stimulus, a typical CDA wave should be observable.

3.2.4 EEG signal pre-processing

As mentioned above, the EEG data were recorded during the task using 64 active electrodes distributed across the scalp according to the extended 10/20 system. An elastic Acti-Cap was positioned on the scalp, with the reference electrode placed on the left earlobe. The high viscosity of the gel used ensured that electrode impedance was kept below 10 KOhm. Signal preprocessing and ERP analyses were performed using BrainVision Analyzer 2.1 (*Brain Products GmbH, Gilching, Germany*).

For the current analysis, we focused on the binocular rivalry experimental condition. Offline re-referencing was applied to the EEG data, using the average activity recorded at the left and right earlobes. The EEG data were segmented into epochs of 1500 ms (−1000/500), time-locked to the first response (i.e., the first reported Gabor patch) following the presentation of each pair of Gabor patches on the screen. Baseline correction was applied, and trials contaminated by ocular artifacts (e.g., blinks or eye movements that exceeded ±50 or ±80 μV, respectively) or other artifacts exceeding ±80 μV were excluded from further analysis. To compute the contralateral waveforms, we averaged the activity recorded by the right hemisphere electrode (PO8) when participants reported seeing the Gabor stimulus on the left side with the activity recorded by the left hemisphere electrode (PO7) when participants reported seeing the stimulus on the right side. The CDA was quantified as the difference in mean amplitude between the contralateral and ipsilateral waveforms within a time window of -700/-200 ms, time-locked to the first response in each trial. Based on visual inspection of the waveforms, a second time window (100/500 ms) was also considered for additional CDA quantification.

3.3 ***Data analysis***

We decided to perform a t-test to observe whether the CDA amplitude was significantly different from zero. In addition, along with the p-value that indicates statistical significance, t-tests provides the Cohen's d. Cohen's d is a value that quantifies the magnitude of the changes that can be observed, in this case, in the CDA amplitude. Thereby, Cohen's d is crucial to understand the practical aspects of the observed statistical significance through the p-value.

Additionally, another advantage of the t-test is an easier comparison to prior and future literature. Using common statistical approaches is important to ensure better and more accurate progress in neuroscientific research, as highlighted by Gambarota and colleagues (2022).

As previously mentioned, we analyzed two CDA time-windows separately, one before the participant gave the answer through the keyboard and one after.

To conclude the description of data analysis, there is another important consideration that has to be made. Initially, 18 participants were recruited for the study. However, we had to exclude 3 participants from the statistical analysis, since the EEG recordings had many artifacts. Other than by blinks or movement artifacts, this phenomenon was also caused by thermal drifts[10], since the research was conducted during summer 2024. Therefore, the sample size included 15 participants (n=15, mean age=24.6, SD=3.48), 14 females and 1 male.

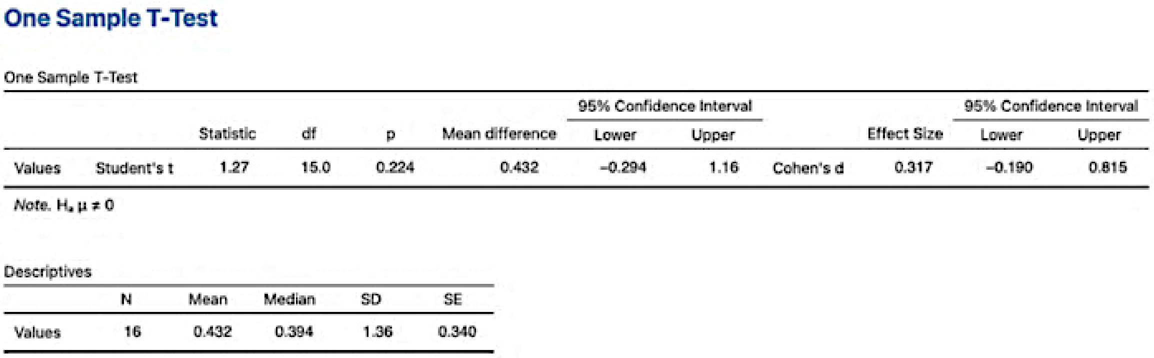
---

[10] Thermal drifts in EEG recordings are frequent artifacts during the summer or during prolonged sessions. Precisely, sweat can alter the conductivity between the skin and electrodes, therefore causing gradual and slow shifts in the base of the EEG signal.

3.4 ***Results***

Firstly, it is fundamental to note that the degrees of freedom (df) are 15. Usually, df are calculated through the formula $df = n - 1$, so the value for n = 15 should be df = 14. However, we had to split the recording of participant 017 in two parts because a problem occurred during the recording phase, leading df to equal 15. More specifically, the server Acti-cap had a malfunctioning due to a problem with the amplifier battery, so we had to start a new recording after approximately two trial blocks from the beginning of the experimental procedure. However, no data was lost so we decided to include this participant in the analysis too.

Furthermore, we will address the results obtained from the two statistical analyses separately. The first result (*see Table 9*) that will be examined is related to the first CDA time-window (-700/-200 ms), which precedes the initial response (e.g., pressing the button "M" or "Z" on the keyboard). As previously noted, after visually inspecting the waveforms, we performed an additional analysis focusing on a different temporal window. More specifically, this second result (*see Table 10*) that will be discussed originates from the second CDA time-window (100/500 ms) which follows the first response on the keyboard given by the observer.



One Sample T-Test

One Sample T-Test

						95% Confidence Interval				95% Confidence Interval	
		Statistic	df	p	Mean difference	Lower	Upper		Effect Size	Lower	Upper
Values	Student's t	1.27	15.0	0.224	0.432	−0.294	1.16	Cohen's d	0.317	−0.190	0.815

Note. H_a $\mu \neq 0$

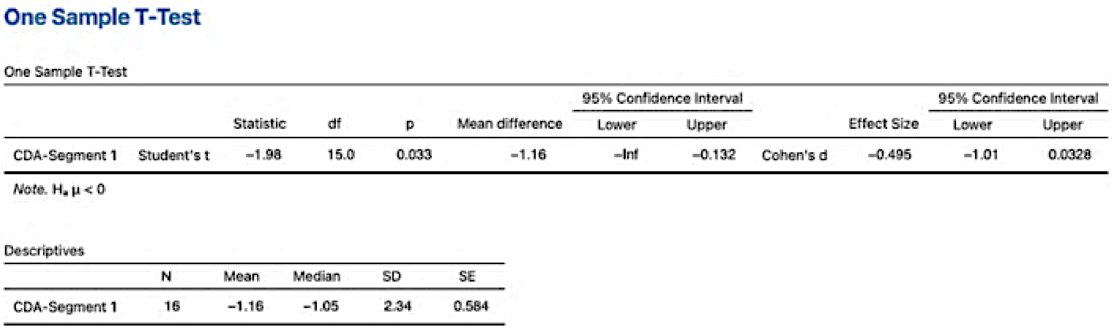
Descriptives

	N	Mean	Median	SD	SE
Values	16	0.432	0.394	1.36	0.340

Table 9. In this table it is possible to observe the results of the statistical analysis made on the first CDA time-window, which precedes the participant's response via keyboard. A one-sample t-test was computed, and the p-value equals 0.224. Since the 95% confidence interval ranges from -0.294 to 1.16, the results are not significant. To phrase this result differently, the CDA amplitude did not significantly vary before the participants gave their first response.

A one-sample t-test was conducted to assess whether the mean amplitude of the CDA component significantly deviated from zero in the first time-window preceding the initial response. The results indicated that the mean amplitude (M = 0.432, SD = 1.36) was not statistically different from zero, t(15)=1.27, p=0.224. The 95% confidence interval for the mean difference ranged from -0.294 to 1.16, suggesting that while there is a slight positive mean amplitude, it is not significant.

The effect size, represented by Cohen's d, was 0.317, with a 95% confidence interval from -0.190 to 0.815, indicating a small to medium effect that is not statistically significant. These results suggest that in this time-window, the CDA component does not show a significant deviation from baseline levels.



One Sample T-Test

One Sample T-Test

						95% Confidence Interval				95% Confidence Interval	
		Statistic	df	p	Mean difference	Lower	Upper		Effect Size	Lower	Upper
CDA-Segment 1	Student's t	−1.98	15.0	0.033	−1.16	−Inf	−0.132	Cohen's d	−0.495	−1.01	0.0328

Note. H_a μ < 0

Descriptives

	N	Mean	Median	SD	SE
CDA-Segment 1	16	−1.16	−1.05	2.34	0.584

Table 10. In this table it is possible to observe the results of the statistical analysis made on the second CDA time-window, which follows the participant's response. A one-sample t-test was computed, and the p-value equals 0.033. Since the 95% confidence interval ranges from -inf to -0.132, the results can be considered significant. To phrase this result differently, the CDA amplitude significantly varied right after the participants gave their first response.

A one-sample t-test was conducted to assess whether the mean amplitude of the CDA component was significantly below zero during the second time-window following the first response. The results indicated a significant negative mean amplitude (M = -1.16, SD = 2.34), t(15)=−1.98, p=0.033. The 95% confidence interval for the mean difference ranged from -inf to -0.132, indicating a reliable negative deviation from zero.

The effect size, measured by Cohen's d, was -0.495 with a confidence interval from -1.01 to 0.033, suggesting a medium effect size in the negative direction.

Additionally, in *figure 11* it is possible to observe the CDA waveform recorded at the electrode sites PO7/PO8. The yellow line represents the CDA, which is our specific ERP component of interest. The dashed red line shows the baseline (0 μV), while the dashed black line represents the response onset, which corresponds to the specific time when the participant provided their first response to the binocular rivalry task.

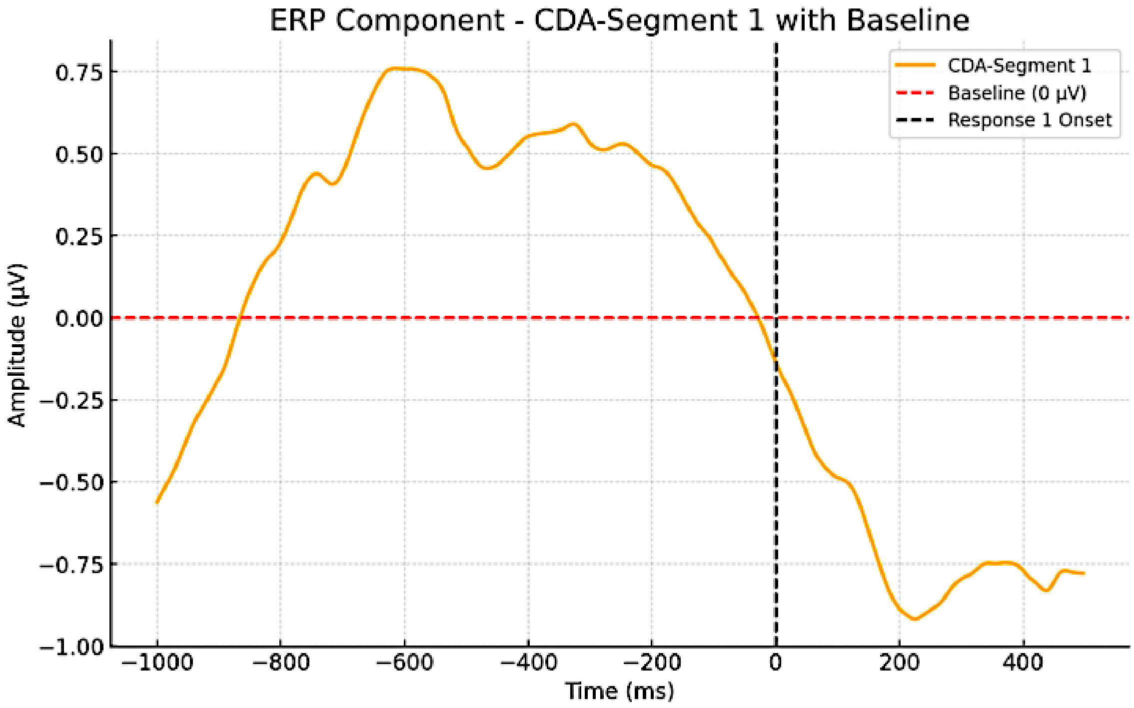


Figure 11. This chart gives a visual representation of how the CDA amplitude varies from -1000 ms before the participants' answers up to 400 ms after. The y axis represents the CDA amplitude, measured in µV, whilst the x axis represents the time, measured in ms.

Finally, it is essential to sum up the results obtained from the statistical analysis.

The findings appear to be quite intriguing and, therefore, they may serve as a foundation to provide additional evidence to the ongoing debate regarding the nature of the representation that VWM holds.

More specifically, the analysis reveals that the CDA is not statistically different from zero until the participant provides a response by pressing the button on the keyboard. This finding potentially favors the initial hypothesis H0, which posits that both conscious and unconscious stimuli could be represented in VWM (then leading to a CDA that does not differ from zero).

Interestingly, the observation of a significant CDA immediately after the participant submitted a response was a result that we did not expect nor anticipate to find. This finding may represent the “working” aspect of VWM, highlighting its ability to maintain relevant information for task performance. In this scenario, the unconscious representation is discarded in favor of the conscious representation, that is needed by the participant to provide a response.

CHAPTER 4 – DISCUSSION

4.1 ***Discussion***

This thesis constitutes an attempt on clarifying whether VWM can maintain unconscious information, thereby contributing additional evidence to the ongoing debate regarding the nature of representations in VWM. The discrete slots model has been consistently validated through several empirical studies (e.g., Luck and Vogel, 1997; Zhang and Luck, 2008; Luck and Vogel, 2013). Consequently, is it possible to conclude that representations in VWM are, at least, discrete. Furthermore, it has been emphasized how these finite number of representations can accurately be reflected by the CDA component. More precisely, the CDA is regarded as an index of VWM capacity, since its amplitude is positively correlated with the number of items maintained in VWM (Luria et al., 2016). Nevertheless, there is an important aspect of how people hold data in VWM that remains unanswered. Specifically, this aspect pertains the possibility of maintaining even unconscious information during WM processes. This phenomenon has caused consistent debate in literature, leading to the theorization of a vastity of ToCs that try to account for all aspects of conscious processing. However, most ToCs only focus on attempting to find the NCCs that could support abstract theorizations regarding what consciousness is and how it can be defined. By contrast, other theories simply try to account for different aspects that are strictly related to consciousness. Notably, several ToCs (Baars, 1998; Tononi, 2004; Lamme, 2006) have been highly influential on these topics. However, Dehaene and Naccache's GNWT (2001) gives a satisfactory framework for consciousness, stating that cognitive processes can occur without conscious representations of the information. According to this theory, conscious processing is possible due to the role of the global neuronal workspace, which enables the rise of a conscious experience through a widespread activation of several interconnected brain regions (Dehaene and Naccache, 2001). Notably, Baars (1997) specifically mentions the WM role stating that WM has some aspects outside the focus of attention (Baars, 1997). Although having many diverse ToCs can be beneficial for research, there are still several aspects that remain unclear, such as the relationship between WM or, more specifically, VWM, and consciousness.

The purpose of the present research is to specifically target the aforementioned debate, trying to give an empirically demonstrated perspective on this topic.

The experimental procedure that has been implemented provides an interesting opportunity to observe unconscious processing of stimuli. More precisely, integrating a binocular rivalry paradigm gives the opportunity to specifically target the potential unconscious maintenance of information in VWM processes (Tsuchiya et al., 2015). Moreover, we chose to use a stereoscopic mirror for creating a binocular rivalry condition because of its compatibility with EEG recordings. In the experiment, we tested the hypothesis that visual working memory (VWM) can maintain unconscious information. Specifically, we expected the CDA (Contralateral Delay Activity) to approach zero if both unconscious and conscious stimuli were being processed. This is because the CDA is a difference wave, calculated by subtracting the ipsilateral activity from the contralateral activity, meaning that equal processing of stimuli in both hemispheres would result in minimal CDA amplitude (Luria et al., 2016).

The experimental paradigm involved trials in which the participant had to indicate the orientation of the bars on the screen using a button on the keyboard. The stimuli consisted of black and white striped Gabor patches that were presented within a circular frame, to facilitate better vergence (Carmel et al., 2010). We chose simple stimuli and not complex ones such as faces or houses because examining how accuracy varies with complexity was not a variable of interest. Regarding accuracy, there is actually no accurate response given via keyboard since, during these types of experimental paradigm, the percept is subjective and therefore alternates in a subjective manner. Consequently, we focused on examining whether CDA amplitude could vary across different trials and, moreover, could reach a value of 0 when both stimuli were presumed to be unconsciously perceived by the participant. As a matter of fact, the ERP analysis corroborates our initial hypothesis. More specifically, the CDA equaled 0 prior to the participant's response. Whilst some may argue that an absence of the CDA could mean that the individual was not paying attention, this seems highly unlikely, considering that this phenomenon could be observed in all trials, for the whole duration of the experiment. Consequently, our initial hypothesis appears to be the most plausible explanation for this result.

In addition to this evidence, we found another result that we did not anticipate. More specifically, the CDA amplitude indicates that one item is being held in VWM, immediately after the observer gives the response via keyboard. There is not a single possible interpretation of this phenomenon. However, according to our opinion, this finding is quite representative of the "working" aspect of memory, as previously highlighted in Chapter 3. Tsubomi and colleagues (2013) have demonstrated that VWM does not merely operate to retain stimuli that disappear from the screen. By contrast, VWM can modulate the conscious representations of stimuli that remain present on the monitor, further highlighting how VWM operates during online processing and does not reflect only "memory" processes per se. To further support this assertion, it is worth mentioning that stimuli did not disappear from the monitor after the observer gave the answer in our experiment. In fact, every presentation of stimulus pairs had a precise duration of several seconds, facilitating binocular rivalry processes and allowing perceptual alternation multiple times per stimulus pair. Nevertheless, Drew and colleagues (2012) have also demonstrated how CDA is a reliable index of VWM even when the quantity of items that need to be held in WM changes across time. More precisely, the authors observed that the CDA amplitude decreased when one representation of the stimuli was lost (Drew et al., 2012). Thus, we can probably assume that this phenomenon may also apply to our experiment. It is reasonable to infer that the increase in CDA amplitude means that the irrelevant representation (i.e., the unconscious one) is lost because holding it in VWM does not serve any purpose for the execution of the task. Therefore, the CDA value is no longer zero because there are not two very similar representations in VWM (i.e., the unconscious and conscious one), but there is only one stimulus representation, which is the conscious one. Additionally, this phenomenon could be interpreted as further evidence of the results obtained by Luria and Balaban (2019b). As it was mentioned above, CDA can accurately reflect resetting processes in VWM, meaning its amplitude can reflect the process of discarding a representation to create a new one (Balaban and Luria, 2019b).

In summary, our findings highly suggest that VWM allows the maintenance of unconscious information. Furthermore, we provide additional evidence that the CDA is a fundamental ERP that can yield significant insights about VWM processes.

4.2 ***Limitations and prospects***

This study integrates two prominent techniques currently available in neuroscientific research: EEG and binocular rivalry. Despite the fact that this study implemented an innovative experimental paradigm, it is not exempt from limitations. As it is often the case, the sample size (n=15) is adequate, but it should be greater to draw more robust conclusions. Nonetheless, considering that this study yielded some promising results, it will be probably extended to provide further evidence on the main topic.

One of the most impactful limitations was that this approach can have a high margin for error. More precisely, as claimed by Carmel et al., (2010), using a mirror stereoscope implies that it has to be adjusted for each participant. Some participants required a relatively straightforward set up, while others found more difficulty in perceiving a single Gabor patch. In addition, adjusting the mirror's position according to the participant's feedback was not always a straightforward process, sometimes resulting in a variable time loss. Notably, it would be beneficial to incorporate a questionnaire regarding discomfort and possible distraction, at the end of the trials. The questionnaire should be filed by the experimenter, that could ask the questions directly to the participant and therefore requiring verbal responses. This approach is preferable since it would be challenging for the observer to answer directly on the keyboard when engaged with the mirror stereoscope. The questionnaire could include items such as "How much discomfort did you experience when viewing the stimuli?" or, for example, "Did you feel distracted during the experimental task?" and ratings should be expressed in a Likert scale. By including this survey, it would be possible to enable better control on the results and to exclude from the statistical analysis the participants who reported being excessively distracted. Moreover, it could provide helpful insights on how to further improve the experimental paradigm.

Another aspect to consider is the equipment that was utilized. Several researchers, including Carmel et al., (2010), recommend placing a piece of cardboard in the middle of the mirror stereoscope to ensure that the observer views the two presented images separately. In our experiment, the piece of cardboard was attached to the mirror with tape and required it to be adjusted with every participant. Therefore, when organizing an experimental setting in the future, it would be better to find a more stable solution. For example, an optimal idea would be to create a 3D-printed structure designed to securely hold the chin guard, the mirror stereoscope and the piece of cardboard, providing more stability for the participant.

Furthermore, the experimental procedure could be refined for future research. As previously suggested by Balaban and Luria (2019), it would be beneficial to include a blank screen for 1 to 2 seconds following each stimulus presentation. By doing so, the participant is allowed to rest the eyes and, therefore, the need for breaks in between trial blocks should diminish. This fact could have an overall effect on the observers subjective feeling of tiredness at the end of the experiment, further reducing the presence of alpha waves on the EEG recording.

Another methodological aspect that needs to be carefully discussed is the lack of a possibility to declare if the viewed item was mixed (i.e., "mixed percept"). It is indeed well known in literature that binocular rivalry paradigms often cause the observer to see an image that is a mix of the two stimuli that are being presented (List and colleagues, 2011). During our research, participants frequently reported to see a mixture of horizontal and vertical bars. Including the possibility of classifying mixed percepts could also enable the investigation on how the CDA amplitude varies in this specific condition, and thereby comparing it to the observed results in our experiment. More specifically, it would be intriguing to assess whether the CDA amplitude remains consistent when the two items are processed simultaneously, one unconsciously and the other consciously. By contrast, the CDA elicited by a mixed percept could be similar to the one observed when maintaining just one image in VWM. To our knowledge, there are no current studies that tried to investigate this phenomenon. Thus, it would be interesting to replicate our experimental paradigm with the suggested modifications. For the present thesis, we tried to avoid confusion by taking into account only the first participants' responses to each trial.

To sum up limitations and prospects, future research should aim to replicate this experimental paradigm to enhance validity, ideally with a greater sample size. While the methodology is innovative, it presents some consistent limitations due to the nature of studying unconscious processing. To investigate these processes effectively, it is crucial to incorporate techniques such as binocular rivalry that do not raise participants' awareness.

CONCLUSIONS

This thesis aims to demonstrate the role of unconscious representations in VWM processes whilst highlighting a new experimental methodology that could be used for future research. Investigating unconscious processing of information is crucial for neuroscience, since the debate regarding whether the brain can process aspects outside the focus of consciousness has been continuing for decades.

Currently, there are no studies in literature that try to address unconscious VWM in the same manner as we did in this study. Our findings, which resulted from combining EEG and binocular rivalry techniques, suggest that unconscious visual representation of stimuli in working memory can occur. To sum up our results, we confirmed our initial hypothesis, which posited that VWM representations could be unconscious, provided the CDA was equal to zero. In addition, a peculiar finding emerged from additional and exploratory ERP analysis. More precisely, we found that CDA accurately reflected the number of representations held in VWM only after the participant responded by pressing the keyboard buttons. This observation suggests that VWM can discard representations that are no longer useful for task completion, therefore emphasizing the "working" component of WM. Moreover, our results align consistently with the GNWT (Dehaene and Naccache, 2001) which tries to elevate the theorization from Baars (1998) towards a more neuroscientific approach, incorporating also NCCs. As a matter of fact, the GNWT postulates that some aspects of cognitive processing can also happen outside of consciousness. That appears, indeed, to be the case with VWM, as Baars (1997) initially hypothesized.

In conclusion, our research provides supporting evidence in favor of the unconscious nature of VWM representations. Notably, however, these results should be carefully interpreted and considered, since our sample size is limited and there have been some methodological inconsistencies along the experimental procedure (e.g., long pauses, stability of the equipment, lack of questionnaires). While it would be premature to consider these findings definitive, they represent a valuable starting point towards understanding more and more about consciousness and how cognitive processes can occur outside conscious awareness.

BIBLIOGRAPHY

"American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature". Journal of Clinical Neurophysiology. ***8*** *(2):200–202. April 1991. doi:10.1097/00004691199104000000007. PMID 2050819. S2CID 11857141./1842/36170.*

Andrade, J., 2001. The contribution of working memory to conscious experience. In: Andrade, J. (Ed.), *Working Memory in Perspective. Psychology Press, pp. 60–78. https://doi.org/10.4324/9780203194157*

Arend, A. M., & Zimmer, H. D. (2011). What Does Ipsilateral Delay Activity Reflect? Inferences from Slow Potentials in a Lateralized Visual Working Memory Task. *Journal of Cognitive Neuroscience*, *23*(12), 4048–4056. https://doi.org/10.1162/jocn_a_00068

Atkinson, R. C., & Shiffrin, R. M. (1968). Human Memory: A Proposed System and its Control Processes. In K. W. Spence & J. T. Spence (A c. Di), *Psychology of Learning and Motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Baars, B. J. (1997). In the Theater of Consciousness. *Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195102659.001.1*

Baars, B.J. (1998). A Cognitive Theory of Consciousness. *Cambridge University Press, 1988*

Baars, Bernard J. (2005). [Progress in Brain Research] The Boundaries of Consciousness: Neurobiology and Neuropathology Volume 150 || Global workspace theory of consciousness: toward a cognitive neuroscience of human experience., *(), 45–53. doi:10.1016/S0079-6123(05)50004-9*

Baars, Bernard J. (2015) Consciousness. *Scholarpedia, 10(8):2207.*

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (A c. Di), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Balaban, H., & Luria, R. (2015). The number of objects determines visual working memory capacity allocation for complex items. *NeuroImage*, *119*, 54–62. https://doi.org/10.1016/j.neuroimage.2015.06.051

Balaban, H., & Luria, R. (2017). Neural and Behavioral Evidence for an Online Resetting Process in Visual Working Memory. *The Journal of Neuroscience*, *37*(5), 1225–1239. https://doi.org/10.1523/JNEUROSCI.2789-16.2016

Balaban, H., & Luria, R. (2019a). Using the Contralateral Delay Activity to Study Online Processing of Items Still Within View. In S. Pollmann (A c. Di), *Spatial Learning and Attention Guidance* (Vol. 151, pp. 107–128). Springer US. https://doi.org/10.1007/7657_2019_22

Balaban, H., & Luria, R. (2019b). Using the Contralateral Delay Activity to Study Online Processing of Items Still Within View. In S. Pollmann (A c. Di), *Spatial Learning and Attention Guidance* (Vol. 151, pp. 107–128). Springer US. https://doi.org/10.1007/7657_2019_22

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science*, *321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Bergström, F., & Eriksson, J. (2014). Maintenance of non-consciously presented information engages the prefrontal cortex. *Frontiers in Human Neuroscience, 8. https://doi.org/10.3389/fnhum.2014.00938*

Bergström, F., & Eriksson, J. (2015). The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. *Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01470*

Bergström, F., & Eriksson, J. (2018). Neural Evidence for Non-conscious Working Memory. *Cerebral Cortex, 28(9), 3217–3228. https://doi.org/10.1093/cercor/bhx193*

Blake, R. A Primer on Binocular Rivalry, Including Current Controversies. *Brain and Mind **2**, 5–38 (2001). https://doi.org/10.1023/A:1017925416289*

Breitmeyer, B. G. (2015). Psychophysical "blinding" methods reveal a functional hierarchy of unconscious visual processing. *Consciousness and Cognition: An International Journal, 35, 234–250.*

Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics, 62(8), 1572–1595. https://doi.org/10.3758/BF03212157*

Brisson B, Jolicoeur P. (2007). A psychological refractory period in access to visual short-term memory and the deployment of visual-spatial attention: multitasking processing deficits revealed by event-related potentials. *Psychophysiology. 2007 Mar;44(2):323-33. doi: 10.1111/j.1469-8986.2007.00503.x. PMID: 17343714.*

Brown, R., Lau, H., & LeDoux, J. E. (2019). Understanding the Higher-Order Approach to Consciousness. *Trends in Cognitive Sciences, 23(9), 754–768. https://doi.org/10.1016/j.tics.2019.06.009*

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional Templates in Visual Working Memory. *Journal of Neuroscience, 31(25), 9315–9322. https://doi.org/10.1523/JNEUROSCI.1097-11.2011*

Carmel, D., Arcaro, M., Kastner, S., & Hasson, U. (2010). How to Create and Use Binocular Rivalry. *Journal of Visualized Experiments, 45, 2030. https://doi.org/10.3791/2030*

Chalmers, D. J. (s.d.) (1995). Facing Up to the Problem of Consciousness.

Clifford, C.W.G. (2009). Binocular rivalry. *19(22), 0-0.doi:10.1016/j.cub.2009.09.006*

Cowan, N. (1995). Attention and Memory: An Integrated Framework. *New York: Oxford University Press.*

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

Cowan, Nelson. (2008). Attention and Memory: An Integrated Framework. Attention and Memory: An Integrated Framework. *26. 1-336. 10.1093/acprof:oso/9780195119107.001.0001.*

De Graaf, T. A., Hsieh, P.-J., & Sack, A. T. (2012). The ‘correlates’ in neural correlates of consciousness. *Neuroscience & Biobehavioral Reviews, 36(1), 191–197. https://doi.org/10.1016/j.neubiorev.2011.05.012*

Dehaene, S. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79(1–2), 1–37. https://doi.org/10.1016/S0010-0277(00)00123-2*

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron, 70(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018*

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10(5), 204–211. https://doi.org/10.1016/j.tics.2006.03.007*

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proc. Natl. Acad. Sci. USA.*

Dehaene, S., Sergent, C., & Changeux, J.-P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proceedings of the National Academy of Sciences, 100(14), 8520–8525. https://doi.org/10.1073/pnas.1332574100*

Drew, T., Vogel, E. K. (2005); Repeated masks are less effective. *Journal of Vision 2005;5(8):75. https://doi.org/10.1167/5.8.75.*

Drew, T., Vogel, E. K. (2008); Neural measures of individual differences in selecting and tracking multiple moving objects. *Journal of Neuroscience 2008, 28 (16) 4183-4191; https://doi.org/10.1523/JNEUROSCI.0556-08.2008*

Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition*, *126*(2), 213–223. https://doi.org/10.1016/j.cognition.2012.10.003

Dutta, A., Shah, K., Silvanto, J., & Soto, D. (2014). Neural basis of non-conscious visual working memory. *NeuroImage, 91, 336–343. https://doi.org/10.1016/j.neuroimage.2014.01.016*

Emrich, S. M., Al-Aidroos, N., Pratt, J., & Ferber, S. (2009). Visual Search Elicits the Electrophysiological Marker of Visual Working Memory. *PLoS ONE*, *4*(11), e8042. https://doi.org/10.1371/journal.pone.0008042

Gambarota, F., Tsuchiya, N., Pastore, M., Di Polito, N., & Sessa, P. (2022). Unconscious Visual Working Memory: A critical review and Bayesian meta-analysis. *Neuroscience & Biobehavioral Reviews, 136, 104618. https://doi.org/10.1016/j.neubiorev.2022.104618*

Gold, J. M., Wilk, C. M., McMahon, R. P., Buchanan, R. W., & Luck, S. J. (2003). Working memory for visual features and conjunctions in schizophrenia. *Journal of Abnormal Psychology*, *112*(1), 61–71. *https://doi.org/10.1037/0021-843X.112.1.61*

Gunseli E, Meeter M, Olivers CN. Is a search template an ordinary working memory? Comparing electrophysiological markers of working memory maintenance for visual search and recognition. *Neuropsychologia. 2014 Jul;60:29-38. doi: 10.1016/j.neuropsychologia.2014.05.012. Epub 2014 May 28. PMID: 24878275.*

Hilimire, M. R., Mounts, J. R. W., Parks, N. A., & Corballis, P. M. (2011). Dynamics of target and distractor processing in visual search: Evidence from event-related brain potentials. *Neuroscience Letters*, *495*(3), 196–200. https://doi.org/10.1016/j.neulet.2011.03.064

Hohwy, J., & Seth, A. (2020). Predictive processing as a systematic basis for identifying the neural correlates of consciousness. *Philosophy and the Mind Sciences, 1(II). https://doi.org/10.33735/phimisci.2020.II.64*

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009

Johnson, M. K., McMahon, R. P., Robinson, B. M., Harvey, A. N., Hahn, B., Leonard, C. J., Luck, S. J., & Gold, J. M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology*, *27*(2), 220–229. https://doi.org/10.1037/a0032060

Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T. Brain areas specific for attentional load in a motion-tracking task. *J Cogn Neurosci. 2001 Nov 15;13(8):1048-58. doi: 10.1162/089892901753294347. PMID: 11784443.*

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible ‘invisible’. *Trends in Cognitive Sciences, 9(8), 381–388. https://doi.org/10.1016/j.tics.2005.06.012*

King, J.-R., Pescetelli, N., & Dehaene, S. (2016). Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. *Neuron, 92(5), 1122–1134. https://doi.org/10.1016/j.neuron.2016.10.051*

Klaver, P., Talsma, D., Wijers, A.A., Heinze, H.J., Mulder, G. (1999). An event-related brain potential correlate of visual short-term memory. *Neuroreport 10 (10), 2001–2005 http://www.ncbi.nlm.nih.gov/pubmed/10424664.*

Klein et al. (2020). Explanation in the science of consciousness: From the neural correlates of consciousness (NCCs) to the difference makers of consciousness (DMCs). *Philosophy and the Mind Sciences, 1(II), 4. https://doi.org/10.33735/phimisci.2020.II.60*

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10(11), 494–501. https://doi.org/10.1016/j.tics.2006.09.001*

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23(11), 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X*

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15(8), 365–373. https://doi.org/10.1016/j.tics.2011.05.009*

Lee, E.-Y., Cowan, N., Vogel, E. K., Rolan, T., Valle-Inclan, F., & Hackley, S. A. (2010). Visual working memory deficits in patients with Parkinson's disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain*, *133*(9), 2677–2689. https://doi.org/10.1093/brain/awq197

Liesefeld, A. M., Liesefeld, H. R., & Zimmer, H. D. (2014). Intercommunication Between Prefrontal and Posterior Brain Regions for Protecting Visual Working Memory From Distractor Interference. *Psychological Science*, *25*(2), 325–333. https://doi.org/10.1177/0956797613501170

List, A., Grabowecky, M., Suzuki, S., (2011). Characterizing mixed percepts during binocular rivalry. *Journal of Vision 2011; https://doi.org/10.1167/11.11.302.*

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400. https://doi.org/10.1016/j.tics.2013.06.006

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031

Luria, R., & Vogel, E. K. (2014). Come Together, Right Now: Dynamic Overwriting of an Object's History through Common Fate. *Journal of Cognitive Neuroscience*, *26*(8), 1819–1828. https://doi.org/10.1162/jocn_a_00584

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

Marcus, D. A., & Soso, M. J. (1989). Migraine and Stripe-Induced Visual Discomfort. Archives of Neurology, 46(10), 1129–1132. *https://doi.org/10.1001/archneur.1989.00520460125024*

McCollough, A.W., Machizawa, M.G., Vogel, E.K., (2007). Electrophysiological measures of maintaining representations in visual working memory. Cortex! 43 (1), 77–94 *http://www.ncbi.nlm.nih.gov/pubmed/17334209.*

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, *11*(1), 103–107. https://doi.org/10.1038/nn2024

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Nakano, S., & Ishihara, M. (2020). Working memory can compare two visual items without accessing visual consciousness. *Consciousness and Cognition, 78, 102859. https://doi.org/10.1016/j.concog.2019.102859*

Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2014). Working memory biasing of visual perception without awareness. *Attention, Perception, & Psychophysics, 76(7), 2051–2062. https://doi.org/10.3758/s13414-013-0566-2*

Perron R, Lefebvre C, Robitaille N, Brisson B, Gosselin F, Arguin M, Jolicoeur P. Attentional and anatomical considerations for the representation of simple stimuli in visual short-term memory: evidence from human electrophysiology. *Psychol Res. 2009 Mar;73(2):222-32. doi: 10.1007/s00426-008-0214-y. Epub 2009 Feb 18. PMID: 19224244.*

Persuh, M., LaRock, E., & Berger, J. (2018). Working Memory and Consciousness: The Current State of Play. *Frontiers in Human Neuroscience, 12, 78. https://doi.org/10.3389/fnhum.2018.00078*

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290. https://doi.org/10.3758/BF03203943

Posner, M. I., & Dehaene, S. (1994). Attentional networks. Trends in Neurosciences, 17(2), 75–79.

Predovan, D., Prime, D., Arguin, M., Gosselin, F., Dell'Acqua, R., & Jolicœur, P. (2009). On the representation of words and nonwords in visual short-term memory:

Evidence from human electrophysiology. *Psychophysiology*, *46*(1), 191–199. https://doi.org/10.1111/j.1469-8986.2008.00753.x

Raffone, A., & Wolters, G. (2001). A Cortical Mechanism for Binding in Visual Working Memory. *Journal of Cognitive Neuroscience*, *13*(6), 766–785. https://doi.org/10.1162/08989290152541430

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (s.d.). Temporary Suppression of Visual Processing in an RSVP Task: An Attentional Blink?

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience, 1(3), 165–175. https://doi.org/10.1080/17588921003632529*

Roy Y, Faubert J. Is the Contralateral Delay Activity (CDA) a robust neural correlate for Visual Working Memory (VWM) tasks? A reproducibility study. Psychophysiology. 2023 Feb;60(2):e14180. doi: 10.1111/psyp.14180. Epub 2022 Sep 19. PMID: 36124370; PMCID: PMC10078237.

Ruchkin DS, Johnson R Jr, Canoune H, Ritter W. Short-term memory storage and retention: an event-related brain potential study. Electroencephalogr Clin Neurophysiol. 1990 Nov;76(5):419-39. doi: 10.1016/0013-4694(90)90096-3. Erratum in: Electroencephalogr Clin Neurophysiol 1991 Apr;78(4):324. PMID: 1699736.

Sandberg, K., Bahrami, B., Lindelov, J. K., Overgaard, M., & Rees, G. (2011). The impact of stimulus complexity and frequency swapping on stabilization of binocular rivalry. *Journal of Vision, 11(2), 6–6.* https://doi.org/10.1167/11.2.6

Schmidt T. (2015). Invisible Stimuli, Implicit Thresholds: Why Invisibility Judgments Cannot be Interpreted in Isolation. *Adv Cogn Psychol. 2015 Jun 30;11(2):31-41. doi: 10.5709/acp-0169-3.*

Seth, A. K. (2018). Consciousness: The last 50 years (and the next). *Brain and Neuroscience Advances, 2, 239821281881601. https://doi.org/10.1177/2398212818816019*

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience, 23(7), 439–452. https://doi.org/10.1038/s41583-022-00587-4*

Shepherd, A. (2000). Visual Contrast Processing in Migraine. *Cephalalgia, 20(10), 865 880. https://doi.org/10.1046/j.1468-2982.2000.00119.x*

Song, C., & Yao, H. (2016). Unconscious processing of invisible visual stimuli. *Scientific Reports, 6(1), 38917. https://doi.org/10.1038/srep38917*

Soto, D., Mäntylä, T., & Silvanto, J. (2011). Working memory without consciousness. *Current Biology, 21(22), R912–R913. https://doi.org/10.1016/j.cub.2011.09.049*

Sperling, G. (s.d.). *THE INFORMATION AVAILABLE IN BRIEF VISUAL PRESENTATIONS.*

Stein, T., Kaiser, D., Hesselmann, G. (2016). Can working memory be non-conscious? *Neuroscience of Consciousness*, Volume 2016, Issue 1, 2016, niv011, https://doi.org/10.1093/nc/niv011

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences, 13(7), 310–318. https://doi.org/10.1016/j.tics.2009.04.006*

Taglialatela Scafati, I., 2019. Is there evidence for non-conscious processing in working memory*? [Ph.D. Dissertation, University of Edinburgh]. http://hdl.handle.net*

Taylor, John G. (2010) Mind-body problem: New approaches. *Scholarpedia, 5(10):1580.*

Thiele, J. E., Pratte, M. S., & Rouder, J. N. (2011). On perfect working-memory performance with large numbers of items. *Psychonomic Bulletin & Review*, *18*(5), 958–963. *https://doi.org/10.3758/s13423-011-0108-7*

Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective & Behavioral Neuroscience, 5*(2), 144–155. *https://doi.org/10.3758/CABN.5.2.144*

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences, 10(11), 502–511. https://doi.org/10.1016/j.tics.2006.09.003*

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience, 5(1), 42. https://doi.org/10.1186/1471-2202-5-42*

Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience, 17(7), 450–461. https://doi.org/10.1038/nrn.2016.44*

Tranel, D., & Damasio, A. R. (1985). Knowledge Without Awareness: An Autonomic Index of Facial Recognition by Prosopagnosics. *Science, 228(4706), 1453–1454. https://doi.org/10.1126/science.4012303*

Trübutschek, D., Marti, S., & Dehaene, S. (2019). Temporal-order information can be maintained in non-conscious working memory. *Scientific Reports, 9(1), 6484. https://doi.org/10.1038/s41598-019-42942-z*

Trübutschek, D., Marti, S., Ojeda, A., King, J.-R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *eLife, 6, e23871. https://doi.org/10.7554/eLife.23871*

Trübutschek, D., Marti, S., Ueberschär, H., & Dehaene, S. (2019). Probing the limits of activity-silent non-conscious working memory. *Proceedings of the National Academy of Sciences, 116(28), 14358–14367. https://doi.org/10.1073/pnas.1820730116*

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural Limits to Representing Objects Still within View. *The Journal of Neuroscience*, *33*(19), 8257–8263. https://doi.org/10.1523/JNEUROSCI.5348-12.2013

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8(8), 1096–1101. https://doi.org/10.1038/nn1500*

Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences, 19(12), 757–770. https://doi.org/10.1016/j.tics.2015.10.002*

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory & Cognition*, *33*(2), 213–220. https://doi.org/10.3758/BF03195310

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. *https://doi.org/10.1038/nature02447*

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. *https://doi.org/10.1038/nature04171*

Voytek, B., & Knight, R. T. (2010). Prefrontal cortex and basal ganglia contributions to visual working memory. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 107*(42), 18167–18172. *https://doi.org/10.1073/pnas.1007277107*

Webster's third new international dictionary of the English language, unabridged. *Springfield, Mass:Merriam-Webster,Editor in chief, Philip Babcock Gove and the Merriam-Webster editorial staff. (2002).*

Wiegand I. Finke K. Müller H. J. Töllner T. (2013). Event-related potentials dissociate perceptual from response-related age effects in visual search. *Neurobiology of Aging, 34, 973–985.*

Woodman, G. F., & Arita, J. T. (2011). Direct electrophysiological measurement of attentional templates in visual working memory. *Psychological Science, 22(2), 212–215. https://doi.org/10.1177/0956797610395395*

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking? Journal of Experimental Psychology: Human Perception and Performance, 22(6), 1505–1513.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860

Zhang, W., & Luck, S. J. (2011). The Number and Quality of Representations in Working Memory. *Psychological Science*, *22*(11), 1434–1441. https://doi.org/10.1177/0956797611417006