



Université de Fribourg
Faculté des Lettres et des Sciences Humaines
Département de Psychologie

Investigating the Impact of Semantic Long-Term Memory on Attention-Based Processes During Working Memory Maintenance

Philippe Schneider

(Brügg, BE)

2022

Thèse de doctorat présentée à la Faculté des lettres et
des sciences humaines de l'Université de Fribourg (Suisse)

Approuvé par la Faculté des lettres et des sciences humaines sur proposition de
Professeure Valérie Camos (première rapporteuse),
Professeure assistante Evie Vergauwe (deuxième rapporteuse) et
Docteur Steve Majerus (troisième rapporteur).
Fribourg, le 16 mai 2022. Le Doyen Professeur Dominik Schöbi

Abstract

Working memory is the system dedicated to the maintenance and processing of information at short time scale. Although this system has attracted a lot of research during the last 50 years, its inner functioning is still a matter of debate. One of the most acknowledged aspect of working memory is that attentional processes play an important role in its functioning, especially during maintenance. Indeed, two attention-based processes have been described during working memory maintenance: attentional refreshing and consolidation in working memory. Attentional refreshing is a domain-general maintenance process that helps maintaining information in working memory against temporal decay and interference. It works by focusing central attention on representation held in working memory, which counteract their forgetting. Consolidation in working memory is also an attention-based process, but it is used to transform fleeting iconic memory traces into more stable working memory traces. In this thesis, we investigated whether attentional refreshing and consolidation are influenced by semantic long-term memory factors. In four series of experiments, we gathered evidence against the hypothesis that attentional refreshing functioning relies on semantic long-term memory. In contrast, we gathered evidence that information that already have a better encoding in long-term memory is consolidated more quickly. Implication of these finding in regard of how we conceive these two processes are discussed.

Acknowledgements

First of all, I would like to thank Evie and Valérie for letting me start the adventure that was my PhD. I am grateful for this opportunity, and I cannot think of better mentors than you have been. Thanks to my lab mates Christophe, Fiona, Luisa, Masami, Jonathan, Stéphanie, and Lady and Gilou for the support, laughs and friendship. I would have not made it without you. Thanks also to all the wonderful colleagues I've met at the University of Fribourg and at the University of Geneva. Tanya, Carli, Caro, Steph, Beatrice, Naomi (and all the other I'm forgetting right now): I hope for the brightest future to all of you! Thanks to Hugo, Nathan and Corentin for the climbing sessions, they were fundamental to my mental health. Thanks to my family and friends who supported me since the beginning; To my father, Margie, Helen, and Alain who paved the way. To my mother and her unconditional and ever-loving support; To Rosa and our endless skypes about the struggle of doing a PhD; To my brother who roasted me for doing it; To my grandmother Jacqueline who greatly enhanced my scientific popularization skills; To Lucien, Martin, Lionel and Rémi for listening to my complaints; To Marion and Mephisto for being present and loving, and finally to all the other people which had an impact on my life, big or small: Thank you!

This thesis is dedicated to Catherine and Theophile Schneider.

Ximita ntsengele xi tshemba nkolo - Shangaan proverb

“Only swallow the fruit if you are confident you can swallow the pip as well”

Table of content

Abstract	3
Acknowledgements	4
Introduction	9
Theoretical Part	13
Chapter 1: Working memory and attention from a theoretical point of view.....	15
1.1 What is attention?.....	16
1.2 Attention in theoretical models of working memory.....	20
1.2.1 The multicomponent model.....	20
1.2.2 The embedded-processes model.....	25
1.2.3 The time-based resource sharing model.....	31
1.3 Conclusion about models of Working Memory.....	39
Chapter 2: Attentional processes in working memory maintenance.....	41
2.1 What is an attention-demanding task?	42
2.2 Attentional Refreshing	45
2.2.1 Controlled Refreshing	46
2.2.2 Swift refreshing.....	49
2.2.3 Limitation to attentional refreshing.....	54
2.2.4 Conclusion on attentional refreshing.....	57
2.3 Consolidation in working memory.....	58
2.3.1 Separability between consolidation, encoding and effect of masked presentation	60
2.3.2 Separability between consolidation and attentional refreshing.....	62
2.3.3 Limits to consolidation	65
2.3.4 Conclusion about consolidation in working memory.....	66
2.4 Conclusion about attentional processes in working memory maintenance	68
Chapter 3: the relationships between working memory and long-term memory from an attentional point of view	71
3.1 Effect of long-term memory factors in working memory.....	71
3.2 How to manipulate long-term memory in a working memory setting.....	77
3.3 Attentional refreshing, consolidation, and long-term memory	82
3.3.1 Attentional refreshing and long-term memory	82
3.3.2 Consolidation and long-term memory.....	84

3.4 Conclusion on the relationship between attention, long-term memory and working memory maintenance.....	87
Experimental Part	89
Chapter 4.1: Visuospatial Familiarity Effect on Attentional Working Memory Maintenance.....	91
4.1.1. Attentional refreshing in working memory	92
4.1.2 The role of semantic long-term memory in attentional refreshing.....	94
4.1.3 Differences between verbal and visuospatial domains.....	96
4.1.4 The current study.....	97
4.1.5 Experiment 1A.....	101
4.1.6 Experiment 1B.....	107
4.1.7 Experiments 2A and 2B.....	111
4.1.8 General Discussion.....	119
4.1.9 Conclusion.....	126
4.1.10 Supplementary Material	128
Chapter 4.2: The Influence of Attentional Refreshing on Face Recognition	131
4.2.1 The present study.....	132
4.2.2 Experiment 1.....	133
4.2.3 Experiment 2.....	138
4.2.4 General discussion.....	144
4.2.5 Conclusion.....	146
Chapter 5: Does Visuospatial Familiarity Influence Consolidation Speed?.....	147
5.1 Consolidation in working memory	147
5.2 Consolidation and encoding processes.....	150
5.3 Is consolidation influenced by familiarity?	155
5.4 The present study	158
5.5 Experiment 1	161
5.6 Experiment 2	171
5.7 General Discussion	188
5.8 Conclusion	193
Chapter 6: Is Guided Refreshing about Attentional Refreshing? Reassessment in the Verbal Domain.....	195
6.1 Maintenance in working memory.....	195
6.2 Additional evidence for an attentional refreshing mechanism	197
6.3 The current study	201

6.4 Experiment 1	202
6.5 Experiment 2	209
6.6 Experiment 3	216
6.7 Cross-Experiment Analysis.....	221
6.8 General Discussion	222
6.9 Conclusion	230
Chapter 7: General discussion	233
7.1 The Impact of Semantic Long-Term Memory on Attention-Based Processes in Working Memory.....	233
7.1.1 Semantic Long-Term Memory and Swift Attentional Refreshing functioning.....	234
7.1.2 Guided refreshing paradigm, controlled attentional refreshing and semantic long-term memory	237
7.1.3 Semantic Long-term Memory and Consolidation in Working Memory	239
7.1.4 The Locus of Long-term familiarity Effect in Working Memory.....	241
7.2 Limits, future studies and methodological advancement	245
7.2.1 Limits to our visuospatial familiarity manipulation.....	246
7.3 Conclusion	248
Bibliography.....	251

Introduction

This thesis aims at investigating attentional processes in working memory and their link to long-term memory. More precisely, we investigated how attentional processes during maintenance are influenced by semantic long-term memory, mostly in the visuospatial domain. Working memory is defined as the ability of our cognitive system to maintain and process information during short time scales (approximately up to 30 seconds). This system is fundamental to our day-to-day functioning. A classic example of the use of working memory is when somebody gives their phone number, the information has to enter the cognitive system, be maintained and then given back the moment it is typed in the phone. However, this example does not represent the full extent of working memory. In addition to this information maintenance role, working memory has also shown to be at the center of “online” processing (to directly process some kind of information, for example adding two numbers, processing a sentence, sorting some numbers, comparing two images, etc.). Although the first models of short-term memory described it as some kind of passive store that is used as a gatekeeper to the entry of information in long-term memory (e.g., Atkinson & Shiffrin, 1968), more recent models acknowledged the findings that the same processes seem to be used for the maintenance of information and processing of information (Baddeley, 2000; Barrouillet & Camos, 2015, 2021; Conway & Engle, 1994; Cowan, 1988, 1999; Daneman & Carpenter, 1980). To emphasize the active part of this system, the term “working memory” was coined, opposed to “short-term memory” which emphasizes a more passive kind of storage.

Since its inception, working memory has attracted a lot of research. However, its inner functioning is still at the center of heated theoretical debates. One of the long-standing theoretical disputes in the field is the relationship between working memory and long-term memory. Long-term memory is an ensemble of systems that can hold a seemingly infinite quantity of information, which can be brought back into working memory if needed. Models of working memory can be put on a spectrum regarding

the separation between working memory and long-term memory: on one end of the spectrum, working memory is described as a separate system from long-term memory (Baddeley, 2000; Baddeley & Hitch, 1974; Barrouillet & Camos, 2015). On the other end of the spectrum, working memory and long-term memory are considered as a unitary system (Brown, Neath, & Chater, 2007). Between these two antipodal positions, intermediary models suppose that working memory is a subset of long-term memory representations, which are in a temporary state of increased accessibility, and this state is qualitatively distinct from representations that are not currently in working memory (Cowan, 1999; Nee & Jonides, 2013; Oberauer, 2002). However, despite the differences in the way working memory and long-term memory are associated, all models acknowledge that a two-way information channel is fundamental between both constructs. Another important point of agreement between models of working memory is the implication of specific attentional processes, especially during maintenance (Awh, Jonides, & Reuter-Lorenz, 1998; Baddeley, 1993; Barrouillet, Bernardin, & Camos, 2004; Cowan, 1995; Kane, Bleckley, Conway, & Engle, 2001; Oberauer, 2009). Indeed, several recent findings showed that specific attentional resources are used during working memory maintenance. Investigating how attentional processes in working memory are modulated by long-term memory could shed a light on the inner interaction between both constructs. More precisely, in this thesis we will focus on attention-based maintenance and consolidation mechanisms in working memory and how they are influenced by semantic long-term memory factors.

The current thesis is structured as follow: in the first chapter of this thesis, we will present several theoretical models of working memory and how they integrate attention and long-term memory. Then, in chapter 2, we will present more thoroughly two attention-based processes thought to happen during maintenance: attentional refreshing and consolidation in working memory. In chapter 3, we will review the current state of the literature on the link between these attention-based processes and long-term memory. Then, we will present the experimental part of this thesis in

which we investigated the relationships between semantic long-term memory and attentional refreshing (chapter 4.1 and 4.2), consolidation and semantic long-term memory (chapter 5), and a novel paradigm that is thought to directly manipulate attentional refreshing: the guided refreshing paradigm (chapter 6). In the final chapter of this thesis, we will present conclusions on our research topic in light of the experiments we conducted.

Theoretical Part

Chapter 1: Working memory and attention from a theoretical point of view

As noted in the introduction, working memory is a central concept in cognitive psychology. Described as “the hub of human consciousness” (Haberlandt, 1997), it is involved when necessary to maintain information for a short period of time and when online processing is needed. For example, following a cooking recipe, following a discussion, finding the way in a new town necessitates the use of working memory. However, despite the huge amount of research dedicated to working memory, its inner functioning is still up to debate. Several models of working memory presently coexist, each with their own strengths and limitations, and are sometimes mutually exclusive. To show the extent of the disagreement among researchers studying working memory, Cowan (2017) showed that at least 9 different definitions of working memory exist in the literature, depending on the research question and the field of study. Despite this apparent lack of agreement, most researchers studying working memory from a cognitive psychology approach agree on the following definition of working memory: a limited capacity system where information is temporary held and processed for ongoing actions and complex cognitions (Cowan, 2017; Logie, Camos & Cowan, 2021). In spite of the apparent diversity of ways working memory has been theoretically described, some aspects of its functioning are recognized by all models. One aspect of working memory that is shared between all the most influential models of working memory is the fact that working memory and attentional processes are either closely related (Awh, Jonides, & Reuter-Lorenz, 1998; Cowan, 1995; Kane, Bleckley, Conway, & Engle, 2001; Oberauer, 2009) or attention is directly controlled by working memory (Baddeley, 1993; Barrouillet, Bernardin, & Camos, 2004), depending on the theoretical background. However, the central issue is the specific content of the “attention” notion. In general, attention refers to the capacity to selectively concentrate on a specific aspect of the person’s external or internal environment. However, the term can be somewhat misleading, as

many other conceptions of attention and how it operates co-exist, and to a greater extent than the term working memory. In the next part of this chapter, we will define several kinds of attention and how they are related to working memory. Then, in a second part, we will describe several models of working memory and how they include specific types of attention.

1.1 What is attention?

In the same vein as the term “working memory” which can be defined in several different ways (Camos, 2017; Cowan, 2017), attention seems to refer to a range of different theoretical constructs. As a consequence, several taxonomies of attention have been described in the literature. For example, Chun, Golomb and Turk-Browne (2011) proposed a taxonomy of attention based solely on the target of attention. These authors propose a distinction between external and internal attention. External attention would be in charge of selection and modulation of perceptually present information, whereas internal attention would be in charge of selection, modulation and maintenance of internally generated information. For these authors, attention (external and internal) is not a unitary construct, but the result of the interaction of a range of cognitive and perceptual operations. This proposed taxonomy, although helpful to distinguish and organize how to conceptualize attention and to generate future research questions, may be too broad to analyze specific cognitive processes involved in working memory context. Another, and maybe more fruitful, way of thinking about the different kinds of attention is to divide attention not only on the basis of the target of attention, but also on the functional use of attention (e.g., Oberauer, 2019). In the following paragraphs, we will describe a taxonomy of attention that distinguishes different kinds of attention. The distinctions rely on three different dichotomies on which attention can be mapped.

First, as proposed by Chun et al. (2011), a distinction can be made between attention to information that is currently present to the senses (image, sound, etc.) and information that is currently not present (basically, the content of ongoing

thoughts). Attention to currently present information can also be called external attention or perceptual attention. Its role is to focus attention on some features of our environment to be processed (automatically or not). In contrast, attention to thought (i.e., internal attention or central attention) is more closely related to working memory. Indeed, as we will see in a later part of this thesis, all models of working memory include this latter type of attention as an important aspect of working memory, whereas fewer model consider perceptual attention as an integral part of working memory and not as part of a peripheral system to working memory. Although this dichotomy has been described in the literature, recent studies showed that the capacity limit of both types of attention was sensibly similar (Tsubomi, Fukuda, Watanabe, & Vogel, 2013). This is congruent with the idea that the same process is at play in between external and internal attention. However, as this debate is not at the center of the present thesis, we will not elaborate further on this topic and keep the distinction between external and internal attention, at least for clarity purpose.

In addition to the external and internal attention dichotomy, another distinction has been made between automatic and controlled attention (Poltrock, Lansman, & Hunt, 1982; Schneider & Shiffrin, 1977; Weichselgartner & Sperling, 1987). Automatic attention refers to the non-controlled attendance to salient features of a person's environment or inner thought. For example, walking down a street and suddenly and unexpectedly hearing a loud bang on the person's side will attract its attention without inducing a conscious wish to turn towards the direction of the loud sound. Controlled attention, in contrast, refers to the conscious act of focusing attention on a specific aspect of the person's external or internal environment. An example is the person in the middle of a crowded place trying to focus on what a friend is saying in the midst of many other people who are talking. The capacity to filter and understand the voice of the friend without listening to every other ongoing conversation is possible thanks to controlled attention, which allows one to consciously focus attention on the specific discussion. It has been shown that automatic attention is

faster and less influenced by task demands than controlled attention, which is slower and more influenced by task demand.

One last distinction between two kinds of attention found in the literature, which is more theoretical and more directly linked to working memory, is the distinction between attention thought as a resource which is responsible for the capacity limit of working memory (e.g., Barrouillet & Camos, 2015; Bays & Husain, 2008; Case, 1972; Just & Carpenter, 1980) and attention defined as a mechanism for the selection and prioritization of information in working memory (e.g., Chun, Golomb & Turk-Browne, 2011, Oberauer, 2009). This last distinction is fundamental to differentiate several models of working memory. The resource approach considers that attention is a resource used during an attentionally-demanding task. The idea is that a certain amount of attentional resources (usually thought as a continuous quantity) can be allocated to specific processes depending on task demand. The more attention is dedicated to a process, the better the performance of this process would be (Navon & Gopher, 1979). In this approach, the capacity-limit of working memory would be explained by the fact that maintaining information in working memory is attentionally demanding. Thus, the more information held in working memory, the more resource demanding it is, up until attentional resources are depleted and working memory attains its capacity limit. On the other hand, the attention-as-selection considers attention as a mechanism to select and prioritize what should be held in working memory in the aim of a cognitive goal, and discard what is no longer relevant. In this approach of attention, working memory can be thought of as attention to memory representations (Oberauer, 2019).

To summarize, the term "attention" can refer to a range of different cognitive functions. It can refer to processing of external or internal features, can be applied automatically or in a controlled fashion and can be conceptualized as a resource or as a selection mechanism. Some of the dichotomies presented above can be intermixed to define intermediary attentional constructs. For example, we can distinguish

between automatic and controlled external attention. The former is the sudden attraction of attention to some features of the environment (the “loud bang” example above) which will automatically attract attention on some features, whereas the former would be used when looking for something specific in the visual environment (for example, when looking for car keys which are not in their usual spot).

In contrast, some dichotomies presented above can become synonyms, depending on the overall theoretical approach. For example, Oberauer (2019), in their taxonomy, considers that controlled attention is resource-limited and automatic attention is not resource-limited. It is important to note that, in our taxonomy, the different kinds of attention are separated on the basis of their functional utility. This does not mean that each kind of attention described above relies on distinct cognitive processes. It may be that some general processes are shared between different kinds of attention, and that other processes are specific to some kind of attention.

In the next part of this thesis, we will investigate the relationships between certain types of attention and working memory. As we will see, all presented models include attentional processes in the functioning of working memory. However, not all types of attention are considered an integral part of working memory. Most models include some type of internal attention, but fewer include peripheral attention as part of working memory.

1.2 Attention in theoretical models of working memory

Now that we have defined attention, we will turn to its relationships with working memory. To this aim, we will first describe several models of working memory and how they include attention, then we will synthesize the points of agreement and disagreement between models on the role and the definitions of attention. For an ease of interpretation, the terms about attention used in the next part will refer to the taxonomy presented above. The three models that will be described have been selected because each one of them highlights an important aspect of working memory functioning.

1.2.1 The multicomponent model

The multi-component model of working memory is the most cited model of working memory to date. First developed by Baddeley and Hitch (1974), the model evolved during the following fifty years in the light of findings observed in the working memory research field. In the last version of this model (Baddeley, Allen & Hitch, 2011), working memory is comprised of two modality-specific "slave" subcomponents used to store and maintain domain-specific information in addition to a domain-general subcomponent, the episodic buffer, which can hold aggregated information from the two other slave components and from long-term memory. An episodic marker can be added on the representation aggregated by the episodic buffer (Figure 1). The two slave subcomponents and the episodic buffer are under the supervision of the central executive module, which is in charge of coordinating the three subcomponents and updating the content of their stores. All subcomponents are thought to have an internal capacity limit, which depends on the current task at hand and characteristics of the information to hold. In the next paragraphs, we will describe the different subcomponents and their interactions in details from the last version of the model.

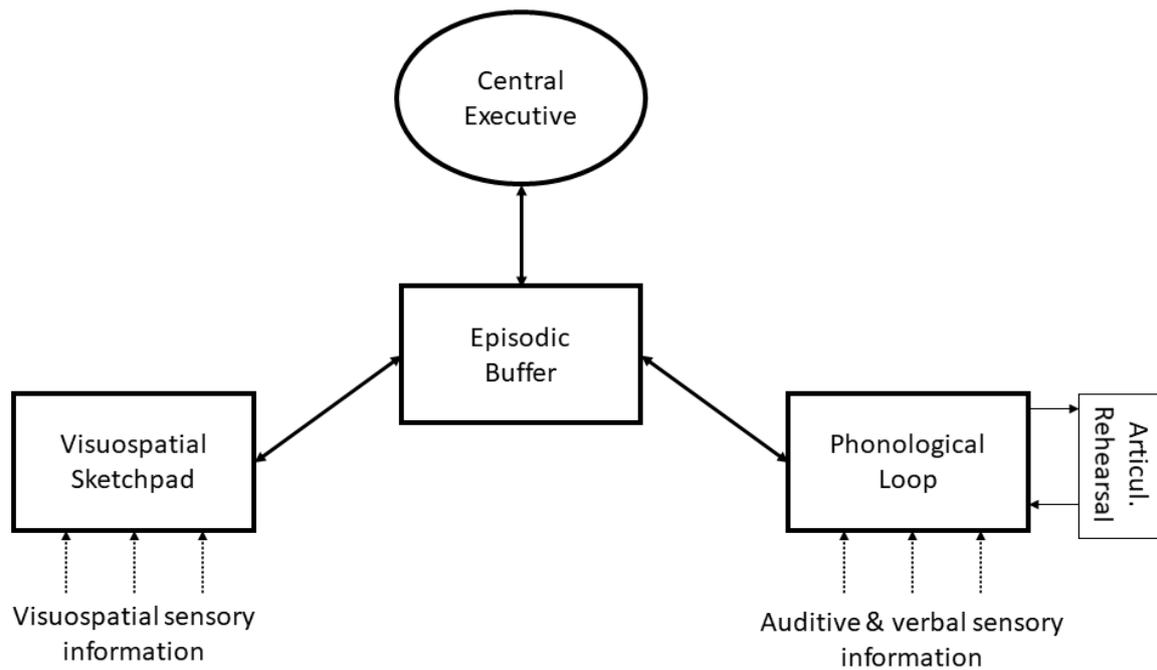


Figure 1: illustration of the multi-component model of Working Memory, adapted from Baddeley, A.D., Hitch, G., & Allen, R. (2021). A multicomponent model of working memory. In: Logie, R.H., Camos, V., & Cowan, N. (eds.) *Working memory: the state of the art*. Oxford (UK): Oxford University Press, p. 31

The first subcomponent, called the “phonological loop”, is used to store auditive and verbal information. It is comprised of two parts: the phonological store, where memory traces are held but are prone to temporal decay and interferences, and the articulatory loop which is used to maintain the memory traces against forgetting. The articulatory loop (also called the rehearsal loop) is thought to work via the vocal or sub-vocal rehearsal of phonological information. The second subcomponent is called the “visuospatial sketchpad”. As its name suggests, this subcomponent is dedicated to the storage and maintenance of visuospatial information. In the same way as the phonological loop, it can be divided in two parts: the store, where visuospatial information is held and prone to temporal decay and interferences, and a maintenance mechanism specific to this store, which protects the information held in the store from being forgotten via temporal decay and interferences. The third subcomponent, called the “episodic buffer”, works a bit differently than the other two subcomponents. It is the last addition to the model (Baddeley, 2000). This subcomponent is a passive system, which hold integrated information coming from

the two other subcomponents and from long-term memory. Contrary to the other two subcomponents, it does not have a specific maintenance mechanism.

The three subcomponents are under the supervision of the central executive module. This component is mainly dedicated to the control of attention on behaviorally and cognitively important information and binding information from different sources (long-term memory, other subcomponents) into a coherent episodic representation to be held in the episodic buffer. In addition, it is also used for the coordination of the slave subcomponents. The authors acknowledge that each component should be divided in set of interacting processes which, put together, explain the capacity limit of each subcomponent. However, the model is admittedly not reached this stage yet, as more work is needed to specify more precisely each of these processes (Baddeley, Hitch & Allen, 2021).

1.2.1.1 Modality-specific stores and attention

The main aspect of this model is its distinction between different domains in working memory. The model supposes that verbal and visuospatial information are treated by separate modules. This supposition relies on experiments using dual-task methodology and neuropsychological evidence. It has been shown that, in a dual task paradigm, performance was lower when same-domain tasks were carried out concurrently compared to when different-domain tasks were carried out concurrently. For example, Baddeley and Hitch (1974) showed that little to no interference was induced by a concurrent non-verbal task on the serial recall of words, compared to a condition where the concurrent task was a verbal one. This is in line with other studies (e.g., den Heyer & Barrett, 1971), which showed the opposite pattern: memory of spatial locations or visual memoranda was more impeded by a concurrent spatial and visual processing task than a concurrent verbal processing task. These two results can be used to support a double dissociation between verbal and visuospatial systems in working memory. These experimental results are substantiated by several neuropsychological cases in which patient had either their verbal component or their

visuospatial component disrupted following lateralized brain damages, while the other component was relatively preserved (Pillon, Bazin, Deweer, Ehrlé, Baulac & Dubois, 1999). Baddeley and colleagues assumed two separate stores for verbal and visuospatial memoranda to account for this double dissociation, thereby also acknowledging that domain-specific attention has to exist and is part of working memory (Hitch, Allen & Baddeley, 2020). Although treated as peripheral attention (in opposition to central attention, more on it later), these domain-specific attentional processes serve a role for the filtering of domain-specific information and for its maintenance.

However, it is important to note two things about the data on which the model is construed: First, there is no pure task. This means that no single cognitive system is responsible for the success of a task. For example, the seemingly simple task of reading a list of words for later oral recall implies the ability, to read the words, encode their phonological and semantic aspect, maintaining them in working memory and then recalling them by overtly saying these words; all of these steps include several motor and pre-motor processes. This cascade of cognitive processes measures not only the verbal aspect of working memory but also many other aspects of the person overall cognitive functioning. Thus, it could be that the double dissociation effectively measured depends on other characteristics of the memoranda and the task than the sole memoranda domain (Morey, 2018). The second aspect is that the authors of this model admittedly investigated the verbal domain much more thoroughly than the visuospatial domain (Baddeley, 2021). This posits the problem that the visuospatial part of the model has been mainly created mirroring the verbal domain, but much less information on its functioning has been directly tested using the multicomponent model framework (but see Darling, Sala & Logie, 2009; Logie, 2011). This has led to recent critics of the multicomponent model which, for example, argued for an absence of a specific visuospatial subcomponent (e.g., Morey, 2018). Nonetheless, the critics do not hinder the multicomponent model from capturing an important aspect of working memory: domain-specific attention has an impact on

working memory functioning. The core theoretical question relies on whether working memory is comprised of several domain-specific stores, or whether it is a domain-general unitary system, the latter being influenced by peripheral domain-specific processes.

1.2.1.2 Attention in the multicomponent model

In this model, the authors consider two aspects of attention (Baddeley, Hitch & Allen, 2021). They distinguish between controlled attention and attentional selection. Controlled attention, on one hand, is defined as resource-dependent and is used for the maintenance of domain-general memory traces and the conscious control of the focus of attention. This type of attention is controlled by the central executive module. Attentional selection, on the other hand, is used to select specific information in the environment and discard distractor. This second type of attention is separated by sensory domain. Returning to our taxonomy presented at the beginning of this chapter, the authors of the multicomponent model aggregated the external and internal distinction of attention with the attention-as-resource and the attention-as-selection definition of attention, respectively. Recently, the model included a new maintenance process controlled by the central executive named attentional refreshing. Attentional refreshing is a domain-general maintenance process that uses central attention to protect working memory representation from temporal decay and interferences (Johnson, 1992; Barrouillet & Camos, 2015; see Camos, Johnson, Loaiza, Portrat, Souza & Vergauwe, 2018 for a review). Although this process was recently added to the multicomponent model of working memory, it has been described before and more thoroughly in other models of working memory, namely the Time-Based Resource Sharing (TBRS). We will thus describe this process in the chapter dedicated to this model. As we will see in later part of this thesis, several models of working memory use this mechanism to explain attentional effect in working memory.

1.2.2 The embedded-processes model

Contrary to the multi-component model of working memory presented in section 1.2.1, the embedded-processes model of working memory (Cowan, 1988, 1999) does not stipulate a direct separation between sensory modalities in working memory. This model postulate that working memory is the part of long-term memory currently activated or under "the focus of attention". This "focus of attention" is an integral part of the model. To follow the classical metaphor, attention can be thought as some kind of "flashlight" that "illuminates" a part of our environment, being internal or external, so it can be processed more efficiently. The "illuminated" part of the environment is called the "focus of attention", which is usually thought to be severely capacity limited. Returning to the separation between long-term memory and working memory, working memory is thus not a short-term store where information can be held for a limited duration separated from long-term memory, but a set of processes and mechanisms that activate representations in long-term memory for their use in ongoing behavioral or cognitive operations. In the next paragraphs, we will describe more precisely the embedded-processes model and its elements.

The model is comprised of four elements, three of which are embedded (or nested) inside one another (thus the name, Figure 2). We will present the elements in nesting order. First, there is long-term memory. This element stores a seemingly unlimited amount of information without duration limitations. Everything that can be represented in our cognitive system is encoded in long-term memory. The second element is the activated part of long-term memory. It is a subset of long-term memory representations that are currently in an enhanced state of activation. This subset of representations can be accessed more easily for ongoing processing and actions, compared to the "non-activated" part of long-term memory. The third element of this model is called the focus of attention. This element is a subset of the activated part of long-term memory representations, which are in an even higher state of accessibility, compared to other representations in the activated part of long-

term memory. The representations inside the focus of attention are the actual content of working memory and are currently used for ongoing behavioral and cognitive operations. The focus of attention is drastically limited in capacity, with a capacity of around four chunks of information. The last element of this model, called the central executive, is similar to the central executive module from the multicomponent model presented above (section 1.2.1; Cowan, Morey & Naveh-Benjamin, 2021). It is used to consciously control the focus of attention and to control voluntary processing. Although the model does not stipulate a fundamental separation between modalities, it acknowledges that differences in how different modalities are actually represented in the cognitive system impact the capacity inside the focus of attention.

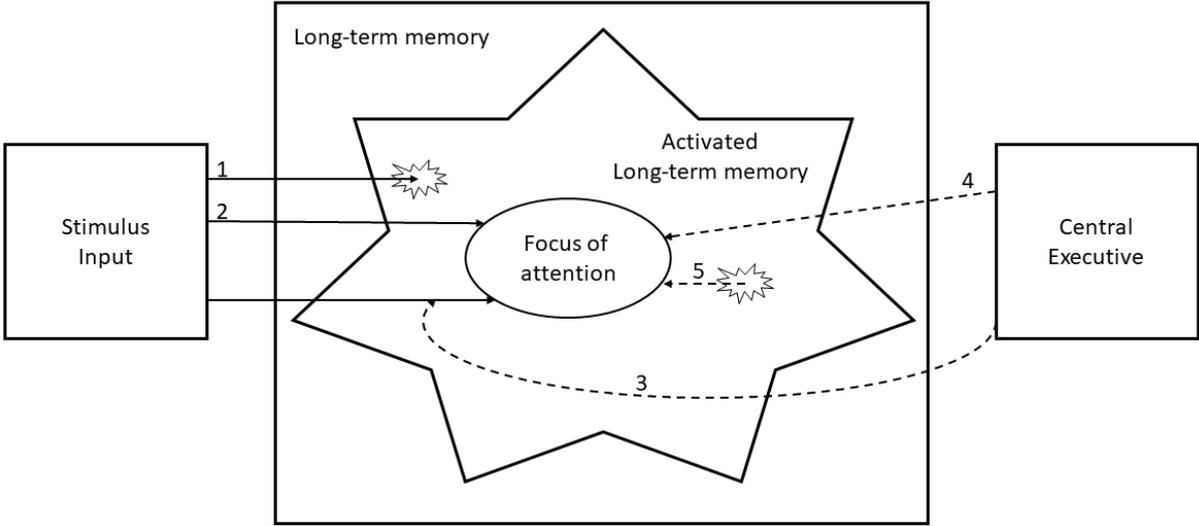


Figure 2: Graphical representation of the embedded-processes model of working memory, adapted from Cowan, N. (2021). An embedded-processes approach to working memory. In R. H. Logie., Camos, V., & Cowan, N (Eds). *Working memory: state of the science* (pp.44-84). Oxford: Oxford University Press.

In this model, the focus of attention is central, as its content is the actual content of working memory. Different ways exist for information to enter the focus of attention or the activated part of long-term memory, some are controlled whereas some others are more automatic. Each numbered path on Figure 2 represents one of these ways, and we will describe all of them. First, information that already has a representation in long-term memory can easily be reactivated when being

perceptually attended (path 1). The representation is simply reactivated into the focus of attention. Secondly, stimulus that have physically changed need to enter the focus of attention to be updated, this also relies on a controlled process, because the changed stimuli also need to be perceptually attended to be processed (path 2). Thirdly, representations can enter the focus of attention when external objects are consciously and deliberately attended (path 3). Fourthly, representations held in long-term memory can be consciously put into the focus of attention, depending on task demand (path 4). Finally, representations in long-term memory can automatically enter the focus of attention if they have a strong enough link with other representations already activated in the focus of attention (path 5).

Since this model considers working memory as the “front end of long-term memory learning” (Cowan, Morey & Naveh-Benjamin, 2021, p.51), the authors included a rapid learning mechanism to accommodate the functioning of working memory. Following their example, if you want to maintain the list of digits “8-5-8”, activating the representation of the digits “5” and “8” is not enough, because the order information would not be encoded. To maintain this list of digits, participants have to activate three serial positions and assign a number to each one, which requires a fast-learning mechanism. The authors argue that this mechanism is identical to the process creating stable representations in long-term memory, although the information is not strengthened enough during its use to be consolidated in long-term memory.

1.2.2.1 Capacity limit in the focus of attention

One important aspect of the embedded-process model is the emphasize put on the capacity limit of the focus of attention. This was due to the interest of the author of the embedded-processes model on individual and developmental differences in working memory and its impact on general cognitive tasks. Indeed, it has been repeatedly shown that working memory capacity is a good predictor of scholastic and intelligence tests (Kail, 2007; Krumm, Ziegler, & Buehner, 2008). In the embedded-

processes model, the capacity limit of working memory is directly linked to the capacity of the focus of attention. This has been substantiated by findings, which showed that it is specifically attentional processes in working memory that are predictive of performance in intelligence and scholastic tests, and not verbal component of the processing (Cowan, Elliott, Saults, Morey, Mattox, Hismjatullina & Conway, 2005; Unsworth & Engle, 2007). Using different methodologies, several studies evaluated the mean capacity of the focus of attention to around four items in healthy young adults (see Cowan, 2010 for a review). This capacity varies from person to person and follows a specific developmental trajectory. Although the embedded-processes model is called a unitary model because it does not suppose a separation between domains in working memory, it acknowledges that memoranda specificities (including the domain of the memoranda) influence the capacity of working memory. For example, higher-level organizational principles can be used to organize information in a more efficient way. The classical example of this is that one can hold approximately 7 separate letters in serial order when they are not related to each other (e.g., "D-T-K-B-S-J-R", Miller, 1956), but a much higher number of letters if aggregated together using long-term memory (e.g., "CIA-FBI-ONU-OMC-UAE-PVC-SVP", Cowan, 2001; Thalmann, Souza & Oberauer, 2019). This dynamic chunking shows that working memory can use peripheral systems to working memory (here, long-term memory) to enhance its capacity. It also shows that the capacity in the focus of attention is not due to a fixed number of slots being filled (e.g., Zhang & Luck, 2008), but that representation in working memory can be dynamically transformed to enhance working memory capacity. Although the authors of the embedded-process model of working memory only recently discarded the view that working memory is limited by a fixed number of slots, they still argue that an overall capacity limit is present inside the focus of attention. On one hand, this capacity limit depends on the characteristic of the memoranda, but, on the other hand, also depends on an overall capacity limits which intrinsically varies from individual to individual.

1.2.2.2 Attention in the embedded-processes model

This model separates at least two notions: controlled attention, exercised by the central executive, and automatic attention, which automatically attracts the focus of attention on specific internal representations. Controlled attention is used to actively control the content of the focus of attention (i.e., updating the content of the focus, inhibiting distractors, maintaining the memory traces active inside the focus of attention), whereas automatic attention seems to be useful for automatically linking information from long-term memory to the content of working memory and thus entering the focus of attention. The authors of the model (Cowan, Morey & Naveh-Benjamin, 2021) consider that the central executive uses the attentional refreshing mechanism to maintain traces in the focus of attention. This attention-based maintenance was later included in the multicomponent model of working memory, as evidence accumulated for its existence, but was first proposed in the embedded-process model of working memory. In addition, this mechanism would also explain the capacity limit of the focus of attention: attentional refreshing is time consuming, and it is not clear if more than one (or a few) representations can be refreshed at the same time. Since the authors acknowledge temporal decay and interferences as a source of forgetting in working memory, and this forgetting is counteracted via attentional refreshing, which can only work on a restricted number of representations at the same time. Only a certain amount of information can be refreshed efficiently until recall, the other representations would be forgotten after the memory traces has been deactivated due to temporal decay and interferences.

To summarize, the embedded-processes model of working memory relies on the assumption that working memory is not comprised of a separate store from long-term memory but is the activated part of long-term memory. This supposes that there is a rapid learning mechanism in long-term memory, otherwise one could not hold information in working memory seen for the first time. It also supposes that attention works as some kind of flashlight that "illuminates" the part of long-term

memory currently needed for action and cognition. The authors of the model also assume that attentional processes, such as refreshing, are resource demanding. They explicitly indicate that the actual resource used by refreshing is the time spent to undergo attentional refreshing. We will see the same idea of time as a resource in the next part of this thesis, where we will describe the time-based resource sharing model of working memory (Camos, Bernardin & Barrouillet, 2004; Barrouillet & Camos 2015, 2021).

1.2.3 The time-based resource sharing model

The time-based resource sharing (TBRS) model of working memory (Barrouillet & Camos, 2001; Barrouillet, Bernardin & Camos, 2004; Barrouillet & Camos, 2015) is based on the idea that working memory is, by essence, a “representational medium, where representations are constructed, elaborated and modified” (Barrouillet & Camos, 2021, p.85). In this approach, working memory is defined as a separate store from long-term memory. However, contrary to the multi-component model of working memory, it is not a simple store where encoded information can be held for a short while, but a medium (or a space) where representations can be built and manipulated. This model is influenced by research from Towse and Hitch (1995, Towse, Hitch & Hutton, 1998, 2002), which stressed the importance of temporal decay and processing speed on the limits of working memory and by the ACT-R architecture (Anderson, 1983; Anderson, Bothel, Byrne, Douglas, Lebiere & Qin, 2004), which put forward the idea that processing in the cognitive system is fundamentally serial, as we will develop below.

The main idea of this model is that representations are not some kind of “objects” held in a seemingly infinite representational space (i.e., long-term memory) that can be reactivated at will or by environmental cues (like in the embedded-process model) or transferred into a working memory store (as in the multi-component model of working memory), but are basically transient objects constructed “on the spot” for the use in an ongoing behavioral or cognitive goal. In this model, working memory is the “space” where representations are created (Barrouillet & Camos, 2015, 2021).

This model is based on four assumptions. The first assumption is that the storage and the processing functions of working memory rely on the same and limited attentional resource. The second assumption is that attentionally demanding cognitive processes can only be done one at a time because of a central processing bottleneck in working memory. This means that maintenance of representations and processing of other representations cannot be done in parallel, but only sequentially.

The third assumption is that representations in working memory will suffer from temporal decay and interferences when attention is disengaged from these representations. Finally, the last assumption is that representations held in working memory can be protected against temporal decay and interference via a rapid switching between the processing of ongoing information and the maintenance of representations held in working memory. This switching between processes takes place during short pauses between processing episodes, which liberates the attentional bottleneck and lets it refresh degraded working memory representations, via focusing attention on it. Following the four main assumptions presented above, the authors of the TBRS model proposed a specific cognitive architecture for their model (Figure 3).

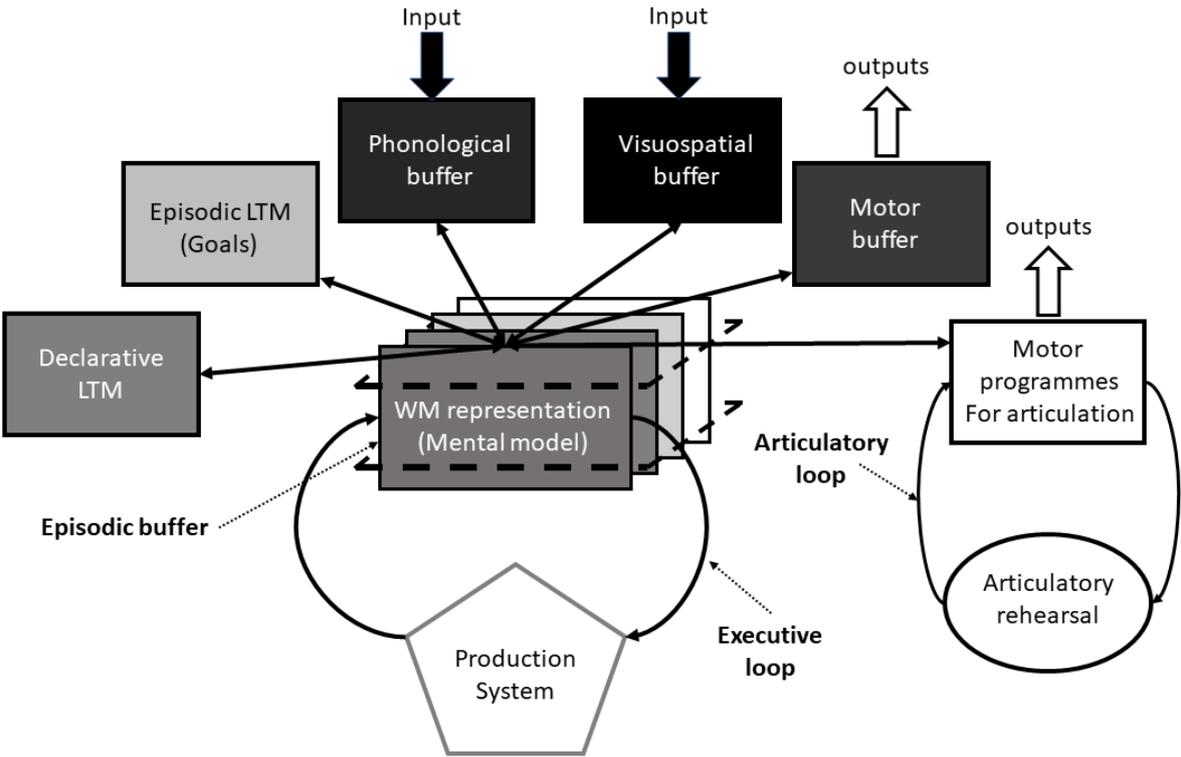


Figure 3: illustration of the TBRS model of working memory, adapted from Barrouillet, P., & Camos, V. (2021). The time-based resource-sharing model of working memory. In R. H. Logie, Camos, V., & Cowan, N (Eds). *Working memory: state of the science* (pp.85-115). Oxford: Oxford University Press.

This model is comprised of a central system, called the episodic buffer, and of several peripheral systems, which include several modality-specific buffers, a production system, and long-term memory. The episodic buffer integrates elements

coming from multiple sensory buffers and long-term memory to create a transient working memory representation (hence the name, in reference to the *episodic buffer* from the multicomponent model of working memory, which serves the same kind of role). The resulting working memory representations are continuously read, transformed, or maintained by the production system hosted in procedural long-term memory. The production system is thought to be comprised of units of procedural knowledge, which follow an "if *condition* then *action*" organization as in the ACT-R model (Anderson, 1983; Anderson et al., 2004). The "if" statement, if filled, will change the representation held in the episodic buffer according to the rule thus activated. As the production system continuously reads from the episodic buffer and in turn influences what is inside the episodic buffer, together they create a so-called "executive loop". This loop is at the basis of working memory dynamic. As said earlier, the working memory representations are created from elements coming from sensory buffer and long-term memory, which are themselves not representations, but elements that can be combined to create a working memory representation. In turn, the representations created in working memory can be sent to peripheral systems, either in declarative long-term memory for long-term storage or to motor buffers for response production.

The main roles of the episodic buffer are the processing of behaviorally important information and the active maintenance of information. However, maintenance of information in working memory does not seem to rely on only one mechanism. The authors of the model defined two maintenance mechanisms. First, there is attentional refreshing. Attentional refreshing is a proposed domain-general and attention-based mechanism that is used for the maintenance of information in working memory. This process is thought to work via focusing central attentional resource on information held in working memory to protect it against temporal decay and representational interferences. This effectively maintains the memory traces held in working memory until they are needed for action. In other words, it is the act of very briefly thinking about what is currently present in working memory to counteract its gradual

disappearance from working memory. Although this mechanism has been added to other models of working memory, research stemming from the TBRS model investigated it more thoroughly. This mechanism is domain-general and is the basic mechanism of maintenance deployed by the executive loop. Since attentional refreshing works via focusing attention on working memory representations, the executive loop will refresh any representation that will occupy the central bottleneck.

The second maintenance mechanism is called articulatory rehearsal. It is specific to verbal information and needs little to no attentional resource to operate. It is thought to work via another kind of loop: verbal information to be maintained is sent to the motor buffer, which holds a motor program necessary for the language production of this verbal information. This information, held in the motor buffer, is then put back into the sensory buffer by playing the information back to oneself (overtly or covertly), thus creating a loop where the verbal information re-enters cyclically working memory. This verbal maintenance mechanism can be disrupted via articulatory suppression, which is the act of repeating a word aloud to prevent the use of the peripheral motor system used in the articulatory rehearsal loop. The inclusion of the articulatory loop in the TBRS model has been influenced by results coming from the multicomponent model of working memory described in the preceding part of this chapter (chapter 1.2.1). Both maintenance mechanisms are thought to be able to maintain around four representations each. Above this limit, representations suffer too much from temporal decay between refreshing episodes to be effectively maintained. However, both mechanisms can also work in parallel and, thus, enhance working memory capacity. A recent study by Barrouillet, Gorin et Camos (2021) showed that, in a special "maxispan" procedure where participants had to maintain series of letters either via articulatory rehearsal, attentional refreshing, or both at the same time, the participants could strategically use both systems in unison to attain a mean span of around 8 items, which is higher than the mean span of 7 items classically found in simple span procedure with letters as memoranda (Miller, 1956).

1.2.3.1 The trade-off between processing and maintenance

One important aspect of this model is the idea that maintenance and processing share the same attentional resource. This would mean that both functions would trade-off their efficiency depending on which process is prioritized. The impact of attentionally demanding concurrent task on recall performance has been shown in several studies (Anderson, Reder, & Lebiere, 1996; Conway & Engle, 1994; Daneman & Carpenter, 1980). To evaluate more closely the impact of a concurrent task on recall performance, Barrouillet and Camos (2015, 2021, Barrouillet et al. 2007, Barrouillet, Portrat & Camos, 2011) extended and more precisely defined a central concept: the *cognitive load* of a task. Usually defined as the intrinsic difficulty of a cognitive task (e.g., Chandler & Sweller, 1991), the authors of this model proposed another definition of the cognitive load. It is the proportion of time during which a specific task captures central attention, and thus impedes other activities that also require central attention. In the specific case of a complex span paradigm (span tasks where the presentation of to-be recalled items is interspersed by a concurrent processing task), the processing steps of the concurrent task can be equated from trial to trial. This simplification allowed the authors to determine that, in the case of a complex-span task, the cognitive load of a concurrent task could be expressed by the following equation:

$$CL = \frac{N \times a}{T}$$

in which "CL" is the cognitive load of the task, "N" the number of processing episodes following each item presentation, "a" the time during which the task occupies the central bottleneck and "T" the total time in between memoranda presentation.

Numerous studies showed that recall performance in a complex span task were heavily predicted by the cognitive load estimated by this equation (up to 98% of explained variance; Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; Barrouillet, Portrat & Camos, 2011) and that both verbal and visuospatial modalities were identically subject to this equation (Vergauwe, Barrouillet & Camos, 2010;

Vergauwe, Dewaele, Langerock & Barrouillet, 2012). The cognitive load effect has been well replicated, which lend to be considered a category A benchmark effect in working memory (Oberauer et al., 2018).

Analyzing the equation, we can see two main aspects of the cognitive load effect. First, the cognitive load of a task is not an intrinsic property of the task in itself but depends on the overall time people have to execute the task. Second, there is several ways to manipulate the cognitive load of a task. To augment the cognitive load, one can increase the number of processing episodes (higher "N"), enhance the difficulty of the task (higher "a"), or reduce the total time participants have to execute the task (lower "T"). In contrast, if two processing tasks have the same overall cognitive load, but one is easier and with less time to do it (lower "a" and higher "T"), and the other has more time overall and is more difficult (higher "a" and lower "T"), then the impact of both processing episodes should be the same on recall performance. The trade-off between processing and maintenance in working memory has also been evidenced by a recent study from Richmond, Burnett, Morrison and Ball (2021) who showed that, in complex-span tasks, working memory capacity can be predicted by performance on the concurrent processing task.

However, since the main idea of the TBRS model is that processing and maintenance processes are trading-off the same resource, maintenance process should also impact processing efficiency. This is exactly what some researchers found. In tasks in which participants have to maintain memoranda and process information in between memoranda presentation (complex-span task) or in task where all memoranda are presented and are followed by a fixed duration retention interval during which participant have to do a concurrent processing task (Brown-Peterson tasks), reaction times to the processing task are influenced by the number of memoranda held in working memory at the time (i.e., the memory load effect; Jarrold, Tam, Baddeley & Harvey, 2011; Vergauwe, Camos & Barrouillet, 2014; Vergauwe & Cowan, 2014). Indeed, mean processing time to the concurrent task was postponed by

approximately 30-50 milliseconds for each new memorandum held in working memory. This showed that processing a concurrent task is influenced by the maintenance of other information, and thus corroborates the trade-off between processing and maintenance proposed by the TBRS model. We will discuss the cognitive load effect, memory load effect and the trade-off between processing and maintenance process more thoroughly in a later part of this thesis dedicated to attentional refreshing.

1.2.3.2 Attention in the TBRS model

Contrary to both models presented earlier, the TBRS model of working memory explicitly defines attention as a resource which can be used for cognitive processes that happen in the episodic buffer and thus occupy the central bottleneck. Moreover, and in contrast to most other approaches which define attention as a resource, the actual resource used during cognitive processes is explicitly defined: it is the amount of time it takes for a specific process to happen. The episodic buffer can only maintain or process one representation at any moment and cognitive processes are themselves time-consuming. Therefore, actual attentional resource used and shared by both functions of working memory is basically the duration during which they occupy the attentional bottleneck. As we saw earlier, this is exemplified by studies that showed a performance trade-off going both ways between processing and maintenance processes in tasks specifically developed to manipulate both processes orthogonally (Barrouillet et al. 2007; Vergauwe et al., 2014).

Coming back to the taxonomy laid out at the beginning of this chapter, the TBRS model distinguishes between attentional processes that uses central attention and these that use peripheral attention. Central attention is used by processes that need to occupy the executive loop, and thus hinder any process done at the same time (i.e., attention as resource). In contrast, peripheral attention is domain specific, will not occupy (or only to a small degree) the executive loop and will hinder only same-domain processing but not different-modalities processing (e.g., visual attention that

will only disrupt visuospatial process, but not auditive process). The peripheral attention is akin to external and automatic attention, as it requires close to no attentional resources.

To summarize, the TBRS model defines working memory as a domain-general medium separated from long-term memory and used for building, maintaining, and processing multi-domain representations useful for ongoing cognitive and behavioral actions. These functions are supported by an executive loop that continuously reads the content of working memory and updates it as needed for the current goal. In this model, attention is defined as a resource needed by everything that occupies the executive loop. Since the executive loop can only process one representation at the time, it creates a processing bottleneck. Due to the fact that maintenance and processing activities both rely on the executive loop, a trade-off between both function is created: representations held in working memory can only be refreshed between processing episodes. Thus, the authors of the TBRS model developed the notion of *cognitive load* of a concurrent task and showed that the cognitive load equation explains this trade-off very well. This time-based trade-off, in turn, explains the capacity limit of working memory and explicitly shows that the actual resource used by attention is not similar to some kind of liquid that is consumed by cognitive operations. It is the actual time during which a process requires attention to be processed in the executive loop.

1.3 Conclusion about models of Working Memory

Although only a subset of models describing working memory, the three models described in this part of the thesis showed a wide variety of paths to model working memory. Due to the apparently immense complexity of working memory, it is not surprising that models differ on several key assumptions and emphasize different aspects of working memory. Nonetheless, we presented these three models, because each one of them captures a different key element about working memory and its link to attentional processes. The multicomponent model highlights that working memory functioning is influenced by the domain of the information to be processed and/or maintained. Although the multicomponent model is the only presented model that supposes a clear separation between verbal and visuospatial working memory, all other models acknowledge that peripheral domain-specific processes also influence working memory functioning. The embedded-processes model put its emphasis on the capacity limit of the focus of attention (and thus on working memory), with a mean estimate at around four items. This capacity limit is an important predictor of many cognitive, developmental abilities and of interindividual differences, which shows the importance of working memory as a whole and its limitation in day-to-day functioning. Finally, the TBRS model stresses the trade-off between processing and maintenance in working memory. It describes more directly the actual relationship between processing and maintenance in working memory, which seems to rely on the same actual mechanism.

From an attentional point of view, all models presented above distinguish between internal attention and peripheral (external) attention. Internal attention is used for the maintenance of information and the processing of attentionally demanding tasks, whereas peripheral attention is used to filter what is entering working memory. Moreover, all models also consider that attentionally demanding tasks require central-attention as the "demanded" resource in contrast to peripheral attention, which relies on domain-specific processes. It is interesting to note that at

least two models presented above (the embedded-processes and the TBRS models) explain the capacity limit in working memory via the same maintenance process: attentional refreshing. In addition, the multicomponent model also recently included this mechanism to explain attentional maintenance in working memory (Baddeley, Hitch & Allen, 2021). However, the precise inner functioning of attentional maintenance processes in working memory is still a matter of debate. Since attentional refreshing is assumed in almost all general models of working memory, investigating its internal working could inform more generally the relationships between attention and working memory.

In the next chapter of this thesis, we will present more thoroughly the attention-based processes during working memory maintenance. First, we will investigate attentional refreshing, then consolidation in working memory. This second mechanism, proposed by Jolicoeur & Dell'Acqua (1998), has attracted research because it could explain several findings regarding attention during maintenance, which are not directly attributable to attentional refreshing. We will see that some conceptions of both processes include long-term memory as an integral part of their inner functioning. Thus, in the final theoretical chapter of this thesis (chapter 3), we will investigate the relationships between attentional processes in working memory maintenance and long-term memory.

Chapter 2: Attentional processes in working memory maintenance

As we saw in the different models presented in the first chapter, attention is defined in several ways and can refer to a whole range of cognitive processes. From attention as a selection mechanism, which chooses what is entering in working memory, to attention as a resource, from external to internal attention, the use of the term attention is ubiquitous in modern cognitive psychology. Since the models presented above aim at describing working memory, they focus on processes specific to internal processing and most of them do not elaborate much on external attention. These models differ on many aspects. Some models describe working memory as separate from long-term memory (e.g., the multicomponent and TBRS models) whereas other models consider working memory as the activated part of long-term memory (e.g., the embedded-processes model). Some models consider that working memory is a unitary system (e.g., the TBRS and embedded-processes models) whereas other models consider that working memory is comprised of several domain-specific stores or buffers (e.g., the multicomponent model). However, there is some aspects of working memory that are shared by all models.

First, as we saw in the preceding part of this thesis (chapter 1.3), they all argue in favor of specific domain-general attentional processes for the maintenance of information in working memory. This is based on a bundle of evidence showing that attentionally demanding tasks impede recall performance when both memory and processing tasks have to be executed in parallel or in close temporal succession. In the literature, at least two attention-based processes, attentional refreshing and consolidation, have been described to explain the effect of attention on working memory maintenance. Second, all models also include a path for long-term memory and working memory to interact. Although the first models of working memory described it as some kind of gateway for entering long-term memory (Atkinson & Shiffrin, 1968), studies and neuropsychological evidence since then showed that the

relationships between the two constructs are more complex and nuanced. However, for now, we will put aside the relationships between working memory and long-term memory as it is the main subject of chapter 3. In the present chapter of this thesis, we will investigate more thoroughly attentional processes in working memory, especially during maintenance.

2.1 What is an attention-demanding task?

As we saw in the preceding parts, models distinguish between attentionally demanding tasks and tasks that do not require central attention. But what is an attentionally demanding task? Without defining it, we risk falling into a cyclical logic where attentionally demanding tasks are defined as tasks that use central attention and central attention is defined as the resource used during these demanding tasks. In the working memory literature, this central attention has also been coined “executive attention” and is akin to controlled and internal attention in our taxonomy (Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007, Engle, Kane, & Tuholski, 1999), or similar to the attention deployed by the central executive (Baddeley, 2000; Cowan, 1999). Executive attention is thus in opposition to external, domain-specific and perceptual forms of attention. The main bundle of evidence for distinguishing attentionally-demanding from other cognitive processes relies on a differential impact of concurrent processing tasks on recall performance. Since it is widely assumed that domain-general maintenance processes in working memory also need central (or executive) attention (Barrouillet et al., 2004, 2007; Cowan, 1999; Daneman & Carpenter, 1980), we can evaluate the attentional demand of a concurrent task by evaluating the level of interferences induced by a concurrent task on recall performance. As we saw previously, it is the logic on which the cognitive load equations have been elaborated (Barrouillet & Camos, 2015; Barrouillet, Portrat & Camos, 2011). Thus, we can summarize that an attentionally demanding process is a task that reliably induces a cognitive load effect, irrespective of the memoranda used.

In the aim of more precisely defining what type of tasks are attentionally demanding, we can make use of the cognitive load equation. If, in a complex span, we use the same concurrent task in different conditions in which we vary the cognitive load of the concurrent task, it is possible to evaluate the impact of the cognitive load manipulation on recall performance of domain-different memoranda. In that case, because the memoranda and the concurrent task do not share the same domain, an effect of cognitive load manipulation should not be explainable via domain-specific interference. If a variation of cognitive load induces a variation on recall performance, we can assume that the concurrent task is attentionally demanding because both processes competed for the same resource. In contrast, if a variation of the cognitive load does not impact recall performance, there are two possibilities: either the concurrent task does not use central attention, or the memoranda cannot be refreshed. To disentangle these two possibilities, a simple solution is to use the same concurrent task with other memoranda. If the cognitive load effect is still not apparent, then the task is not attentionally demanding. If, in contrast, the cognitive load effect appears with other material, then it is the material from the first experiment that is not refreshable.

Using this logic, it has been established that all attentionally demanding tasks share the same features (see Barrouillet & Camos, 2012 for a review). They require a response selection (e.g., yes/no; up/down) on a semantic task or have to intermix different modalities (and thus occupy the executive loop/focus of attention, depending on the theoretical approach). A domain-specific concurrent task (e.g., in the visuospatial domain, deciding which of the two simultaneously presented lines is the longest) will not use central attention much but will induce interferences for the same-domain memoranda. Indeed, the processing of this task can solely rely on domain-specific perceptual processing (Vergauwe, Barrouillet & Camos, 2009). In contrast, asking for the parity of a digits presented in Arabic format will require executive attention, because the numerosity has to be extracted and then a response selected depending on the parity of the digits. Several studies have shown that parity

judgement relies on central attention (e.g., Vergauwe et al. 2010, 2012). Now that we have defined what is an attentionally demanding task, we will present one of the more investigated processes which requires attention in working memory maintenance: attentional refreshing.

2.2 Attentional Refreshing

We succinctly described attentional refreshing above when presenting the models of working memory, but now we will describe more thoroughly. Attentional refreshing is a proposed maintenance process that uses central attention (i.e., attention controlled by the central executive in the multicomponent model or in the embedded-process model; or occupy the central buffer in the TBRS model) to maintain memory traces in working memory for ongoing behavior and cognition, irrespective of the domain. First theorized by Johnson (1992), this proposed mechanism attracted a lot of research in the last 15 years (see Camos, Johnson, Loaiza, Portrat, Souza & Vergauwe, 2018, for a review). Despite the amount of research dedicated to it, attentional refreshing has remained rather mysterious. Several approaches described it in slightly different but significant ways.

The first debate about the inner working of attentional refreshing relies on its speed and deliberateness. On this basis, two kinds of attentional refreshing have been described. On one side, researchers defined attentional refreshing as a deliberate process where attention would be consciously put on representations held in working memory to counteract their disappearance from working memory due to temporal decay and interference. Said differently, it is the act of actively thinking, “imagining”, about information currently held in working memory (Johnson, 1992; Souza, Rerko & Oberauer, 2015). This process is thought to span several hundred of milliseconds and is a conscious, deliberate act. We will coin the term “controlled refreshing” to describe this form of refreshing. On the other side of the debate, researchers described attentional refreshing as a swift process, lasting around 30 to 50 milliseconds (Camos et al. 2019; Vergauwe, Camos & Barrouillet, 2014). The term “swift refreshing” has been coined to describe this kind of attentional refreshing (Camos et al., 2018). Both conceptions of refreshing were mostly investigated separately, using different experimental designs. The controlled refreshing has been mainly investigated via retro-cue paradigms, whereas swift refreshing has been

mostly studied using complex-span and Brown-Peterson tasks. For now, we will describe further controlled refreshing. Swift refreshing will be investigated in a later section of this chapter.

2.2.1 Controlled Refreshing

Controlled refreshing, as said earlier, is described as a deliberate and conscious act. The main bulk of evidence for its existence has been gathered using a “multiple retro-cues” paradigm (or guided refreshing paradigm). This paradigm has been created on the basis of the “simple” retro-cue paradigm. In a simple retro-cue paradigm, a set of memoranda is presented to the participant for a fixed duration. Then, during a retention interval in between presentation and recall, a cue appears on screen highlighting the location of one of the memoranda (i.e., the memorandum has been *cued*). The classic result of this manipulation is that memoranda that have been cued induce better recall performance than non-cued memoranda (Griffin & Nobre, 2003; Landman, Spekreijse, & Lamme, 2003). This paradigm is a sort of evolution from the classical cueing paradigm (Posner, 1980), in which a cue presented before memoranda presentation will enhance the recall performance of the memoranda that was cued. The interesting aspect of the retro-cue paradigm is that the cue appears after a delay following the memoranda presentation. This ensures that the effect relies on working memory and not on iconic memory, as the delay is classically well over 500ms (which is considered outside of iconic memory range). The multiple retro-cues paradigm is an evolution from the simple retro-cue paradigm, where multiple retro-cues are presented during the retention interval. The main manipulation is that memoranda can be cued 0, 1 or 2 times and that participants are tasked to actively think about the memoranda that have been cued.

In the first experiments using this paradigm (Souza, Rerko & Oberauer, 2014), participants were simultaneously presented six colored dots, evenly distributed around a central point in the middle of the screen, for one second. Participants were tasked to memorize this array of colored dots. Then, after another one-second delay,

four cues appeared sequentially on screen, each for 500 milliseconds. They consisted of arrows that could point in the direction of one of the presented dots each. Participants were tasked to actively think about the colored dot presented to which each cue pointed. After the retro-cueing phase, one of the emplacements where the colored dots were presented was highlighted and participants had to reconstruct the color presented during the study phase at this emplacement, using a color wheel. Results showed that the more the probed emplacement was cued, the better was precision in recall performance (distance in degrees between the presented color and the reconstructed color on the color wheel). The authors (Souza, Rerko & Oberauer, 2014) thus argued that, by thinking about the colors the cues pointed at, participants actively refreshed their representation and protected it against interference¹.

These authors called this multiple retro-cue paradigm the “guided refreshing paradigm” because, in their opinion, one could guide on which representation refreshing would occur more. However, their finding had several limitations: it was limited to the visual domain and item presentation was always simultaneous. Since attentional refreshing is defined as a domain-general mechanism, the effect of the guided refreshing should also be evident in the spatial and verbal domain. In a follow-up study, Souza, Vergauwe and Oberauer (2018) investigated these possible confounds on the effect of the guided refreshing paradigm. The study comprised several experiments. In Experiments 1a and 1b, the researcher manipulated spatial memoranda (positions around a circle) and the type of cues that were used (peripheral or central). In both experiments, they replicated the effect of the number of retro-cues on recall performance. In Experiment 2, they replicated the results from the 2014 study, with the difference that the colored dots were presented sequentially

¹ It is worth noting that, from a theoretical point of view, the authors of this study deny temporal decay as a source of forgetting in working memory. In their view, the disappearance of memory traces from working memory is solely driven by interferences. Thus, they do not mention protection from temporal decay in their conclusion. We do not expand on this theoretical debate, as it is not in the scope of this thesis.

instead of simultaneously. Again, the effect of the multiple retro-cues was evident, with an increase in precision of recall performance for each additional cue pointed to the memoranda. However, in Experiments 3a, 3b and 3c, the authors investigated the guided refreshing with verbal material and found another pattern of results. These three experiments followed the same experimental design with minor changes of some parameters between the three experiments. These experiments in verbal material were designed as follows: six boxes were presented on screen, each at an equal distance from the center of the screen and arranged in circle. Six words appeared sequentially, each in one of the boxes. Participants were tasked to read and memorize each of these words. Then, after a 500-millisecond interval, four cues appeared sequentially on-screen. Each cue consisted of an arrow that could point in the direction of one of the boxes (which stayed on-screen for the whole duration of a trial). After the cueing phase, one of the boxes was highlighted and participants had to recall the word that was presented in this box during the presentation phase.

The three experiments differed as follows: Experiment 3a used an oral recall procedure, whereas Experiment 3b and 3c used a typed recall procedure. The other differing aspect between the three experiments was the presence or absence of articulatory suppression. Experiment 3a and 3b used an articulatory suppression method (repeating "BABABA" out loud throughout the trials) whereas participants could verbally rehearse as they chose in Experiment 3c. Results in the verbal domain were much more ambiguous than in the visuospatial domain. Evidence for an effect of the retro-cue manipulation was ambiguous for Experiments 3a and 3b ($BF_{10} = 1.5$ and 0.44 , respectively) but present for Experiment 3c ($BF_{10} = 7.85$). Since Experiment 3c was the only one where participants were not under articulatory suppression, it may be that participants could have actively rehearsed words that were cued compared to non-cued words, and thus recalled them better. In contrast, when participant could not rehearse the retro-cued words, evidence for an effect of the retro-cue was much more ambiguous. Although the effect of the number of retro-cues become evident when the three experiments are analyzed together

($BF_{10} = 120$), these results nonetheless put into doubt that the multiple retro-cue effect is effectively domain general. If the effect in the verbal domain is driven by verbal rehearsal, this will in turn cast doubt on the direct implication of domain-general attention on the multiple retro-cue effect detected in the visuospatial version of the task. Although it has been shown that general attention is involved during the overall task (Souza & Oberauer, 2017), it is not clear whether the multiple retro-cue effect stems from attentional processes in themselves.

To conclude, some evidence has been gathered for the existence of the controlled refreshing mechanism, but mostly in the visuospatial domain and via the guided refreshing paradigm. However, doubts remain on the actual implication of domain-general attention in this paradigm. In chapter 6 of this thesis, we directly investigated possible alternative explanations for the absence or presence of the effect in the verbal domain.

2.2.2 Swift refreshing

Contrary to controlled refreshing, which is conscious and deliberate, swift refreshing is defined as extremely quick (few tens of milliseconds) and thus outside of a direct conscious control. Proponents of this definition of attentional refreshing describe it as the act of putting representations in the center of the focus of attention (e.g., in the embedded-processes model) or in the episodic buffer (e.g., the TBRS and multicomponent models). Refreshing could occur every time attention is not engaged in a concurrent processing activities. As said in a review about attentional refreshing (Camos et al., 2018), swift refreshing is an “atomic” episode of a larger maintenance process. Several refreshing episodes could happen in short succession in-between processing episodes. The main bulk of evidence regarding swift refreshing has been gathered using complex-span and Brown-Peterson tasks. Both tasks have been developed to investigate maintenance and processing of information in working memory orthogonally.

As we saw in the part dedicated to the TBRS model of working memory, the cognitive load of a task (calculated via the equation $CL = \frac{N \times a}{T}$) is a very good predictor of recall performance in the simplified situation of a complex-span task. This relation between cognitive load and memory task performance is the same for verbal and visuospatial memoranda. The higher the cognitive load of a concurrent task, the worse the recall performance of the memoranda held in working memory (i.e., the cognitive load effect). The authors who developed the cognitive load equation explained this trade-off via the swift refreshing mechanism (Barrouillet et al. 2007; Barrouillet & Camos, 2015, 2021; Vergauwe et al., 2012): refreshing episodes would happen more during the overall retention interval when the cognitive load of a task is low. This would, in turn, enhance maintenance opportunity compared to a situation where the concurrent task has a higher cognitive load. A higher cognitive load would occupy the central bottleneck for longer, thus limiting the possibility to engage in attentional refreshing. To illustrate this point, let's take a hypothetical complex-span task in which the main manipulation is that two concurrent processing tasks (A and B) are orthogonally manipulated. Participants will have to memorize the same set of memoranda, but they will either have task A or task B as the concurrent processing task. Task A takes more time than task B (higher "a") but the overall retention interval in both conditions is the same (Same "T"). Following our equation, this would mean that task A has a higher cognitive load than task B. Since the overall retention interval is kept equal, participants will have more free time during the retention interval of task B to switch from the processing to attentional refreshing. In turn, information held in working memory will be refreshed more frequently than when task A is used as the concurrent processing task. This will result in better recall performance when task B is used, compared to when task A is used.

The cognitive load argument for the existence of swift refreshing relies on the supposition that processing and maintenance of information in working memory competes for the same resource. If this is the case, we should also detect an effect of memoranda maintenance on a concurrent processing task (i.e., a memory load effect)

in addition to the cognitive load effect (impact of a concurrent task on memoranda maintenance). Several studies actually tested this hypothesis and did find an effect of maintenance on a concurrent processing task (Camos, Mora, Oftinger, Mariz-Elsig, Schneider & Vergauwe, 2018; Vergauwe, Camos, & Barrouillet, 2014; Vergauwe & Cowan, 2014). Adapted by Vergauwe, Camos and Barrouillet (2014), most of these results used a Brown-Peterson paradigm (Brown, 1958; Peterson & Peterson, 1959). In this task, a variable number of memoranda is presented sequentially at the beginning of a trial. They all stay on screen for the same duration. Then, after the last memorandum for the trial disappears from screen, a fixed-duration retention interval starts. During this interval, participants need to process a concurrent task. They are tasked to process as many items as possible during this interval. Then, after the retention interval, participants have to recall the memoranda presented at the beginning of the trial (usually in the order they were presented in). In the adapted version of Vergauwe and colleagues (2014), the main manipulation is the number of items presented at the beginning of a trial. For the analysis, the authors only keep trials in which the memory task was successful (correct memoranda recalled in the correct position) to ensure that all memoranda were effectively maintained. The results showed that, for each memorandum held in working memory, reaction times to the processing task were postponed by approximately 30-50 milliseconds. The authors took this postponement as evidence for the trade-off between processing and maintenance. The idea was that all memoranda are refreshed in between processing episode, and thus more memoranda needed more time to all be refreshed compared to a situation where less memoranda are held in working memory. This methodology also helped to evaluate the speed at which refreshing works. If we assume that refreshing is a serial mechanism that will refresh all memoranda held in working memory sequentially, then the 30-50ms postponement can directly be interpreted as the speed it takes to refresh one representation held in working memory. It is important to note here that some specific memoranda did not induce a memory load effect. This poses the question of the limitations of swift attentional

refreshing. We will tackle this important aspect in the next part of this chapter, but for now we will focus on the possible schedule of attentional refreshing.

The idea that attentional refreshing happens in the 30-50 millisecond range relies on the supposition that all memoranda are refreshed sequentially, but in an all-or-nothing manner. However, the schedule on which attentional refreshing is applied to representations in working memory is still a matter of debate. Although the 30-50ms postponement have been replicated numerous times, it is not clear yet what is the real meaning of this postponement. Several schedules have been put proposed (Lemaire, Pageot, Plancher & Portrat, 2018; Oberauer & Lewandowsky, 2011). Refreshing could happen sequentially on all memoranda, starting from the first encoded one to the last encoded one: Alternatively, it could also happen sequentially, but from the least activated memoranda, or from the most activated one. Importantly, refreshing could also be a simultaneous process: more than one representation held in working memory could be refreshed at the same time. If this is the case, it would mean that the 30-50 milliseconds measured in the Brown-Peterson does not reflect one refreshing episode, but a bigger and more general refreshing process.

To tackle this issue, Lemaire et al. (2018) used a computational modelling approach. They based their study on the TBRS model. Using the TBRS*, a computational representation of the TBRS model of working memory developed by Oberauer and Lewandowsky (2011), Lemaire et al. (2018) implemented six possible attentional refreshing schedules and compared their capacity to model already existing data taken from experiments with complex-span tasks. They implemented possible schedules like random order refreshing (the memoranda are refreshed randomly during free time), cumulative refreshing (all memoranda are refreshed in order of presentation, always starting from the first memoranda presented), last-encoded refreshing (the last encoded item is refreshed in priority), below threshold refreshing (all traces which are below maximum activation are refreshed), probabilistic

refreshing (memoranda are more probable to be refreshed if their activation is low), least-activated first refreshing (the least activated memoranda is refreshed in priority) and the expanded attentional focus (up to four memoranda can be refreshed simultaneously).

Their results showed that the refreshing schedule fitting the data the best was the “least activated first” schedule. This schedule is also interesting, because it does not need to invoke some kind of homunculus to explain how it works. It is sufficient to have a process that could evaluate the strength of activation of a trace held in working memory and trigger refreshing when it attains a certain threshold. However, these results do not seem to be definitive. Another study was more congruent with the last-encoded refreshing schedule. Vergauwe, Hardmann, Rouder, Roemer, McAllaster and Cowan (2016) investigated if memoranda are serially refreshed in working memory. The authors reasoned that if it was the case, different memoranda held in working should have a different activation depending on when the item was probed. So, if they manipulated the stimulus onset asynchrony (SOA) between item presentation and a probe in steps of 100 milliseconds, different memoranda should be in a state of heightened activation at a specific time. The reason is that some memoranda have just been refreshed whereas other are at the end of the list, and more activated memoranda should induce faster response times because of their preferential place inside the focus of attention. Interestingly, they found that the delay had no impact on reaction times to the probe. Their results showed that participants responded more quickly only to the last presented memoranda. This can be taken as evidence (although not final either) that only the last presented item in working memory is in a heightened state of activation, and thus have been refreshed.

To summarize, swift attentional refreshing has received a wealth of evidence of its existence. However, one could argue that this evidence only relies on manipulation of the attentional refreshing *opportunity* to be deployed and not on a direct refreshing manipulation. Indeed, complex-span and Brown-Peterson tasks manipulate the

retention interval and the difficulty of the concurrent processing task. Thus, they manipulate the opportunity to engage in attentional refreshing, but not attentional refreshing directly. Although it is a limitation of the complex-span and Brown-Peterson paradigms, the huge amount of research on the swift attentional refreshing using the cognitive load effect, its domain-generalty and its very replicable effects on working memory maintenance all point in the direction of the existence of swift refreshing as a maintenance mechanism in working memory.

2.2.3 Limitation to attentional refreshing

We have already described the evidence for the existence of two kinds of attentional refreshing. Most of the research on refreshing has been done using complex-span and Brown-Peterson tasks that showed a trade-off between processing and maintenance of information in working memory. However, several studies have shown that specific kind of memoranda failed to be influenced by cognitive load and memory load manipulations. This has raised some questions about the limitations of attentional refreshing and on which kinds of representations it can be applied. Ricker and Cowan (2010) showed that specific kinds of memoranda in the visuospatial domain (unconventional characters) were lost at a similar rate, irrespective of the presence of an attentionally demanding concurrent task or not. This was in contrast to other kinds of memoranda (letters) that suffered a lower forgetting rate when the retention interval was unfilled compared to a condition where participants had to execute an attentionally demanding concurrent task during the retention interval. For these authors, this differential impact of attentionally demanding concurrent task on forgetting rate of memoranda was taken as evidence that some kinds of features, at least in the visuospatial domain, could not be refreshed at all. In their conclusion, these authors proposed several possibilities explaining this differential effect between types of memoranda. It may be that some peripheral visual features cannot be refreshed, due to the overall architecture of the visuospatial processing network. This would mean that some features cannot be refreshed at all, under any circumstances.

Another possibility is that only memoranda that already have a stable enough representation in long-term memory can be refreshed, because attentional refreshing relies in part on long-term memory representations to function.

The idea that some types of features could not be refreshed (at least via the swift refreshing mechanism) in working memory have been substantiated by more recent findings. Vergauwe et al. (2014), using a Brown-Peterson task, varied the type of memoranda (visual, verbal, or spatial) and the kind of concurrent task, which demanded either central attention or visuospatial attention only, in seven experiments. As concurrent task, participants had to either judge the parity of numbers presented sequentially (i.e., a parity task) in experiments two to six, which uses central attention, or participants had to judge whether a line fit between two points (i.e., spatial fit judgement task) in experiment 1, or had to judge whether the screen was more red or more blue (i.e., color discrimination task) in experiment 7, both of which rely on peripheral visuospatial attention. In the first five experiments, where memoranda could be spatial location (Experiments 1 and 2), letters (Experiments 3 and 4), or monosyllabic and disyllabic words (Experiment 5), results showed that reaction times to the concurrent task were a function of the number of memoranda held in working memory.

This memory load effect, as we saw earlier, has been taken as evidence for the impact of maintenance on the processing of a concurrent task (and thus of the two-way trade-off between processing and maintenance in working memory). However, Experiments 6 and 7 showed a different picture: when different fonts of the same letter were used as memoranda, and that participants had to serially reconstruct which font was presented at the beginning of the trial, no memory load effect were detected on the reaction time to the concurrent parity task. Since the parity judgement task has been shown to reliably induce a cognitive load effect, these authors concluded that we could not refresh features from the fonts. In line with the

proposition of Ricker and Cowan (2010), fonts are peripheral features to the letters, and it could be that fonts are not very well-represented in long-term memory.

Another study (Ricker & Vergauwe, 2020) showed that some other material could not be refreshed. In four complex-span tasks, they failed to find an effect of cognitive load manipulation on specific visuospatial memoranda. In their study, the authors used continuous position around a circle (Experiments 1a and 1b) or canonical position around a circle (Experiments 2a and 2b) as memoranda and manipulated the type of concurrent tasks (auditory tone task in Experiments 1a and 2a, and parity judgement task in Experiments 1b and 2b) and their cognitive load (inside each experiment). The four experiments failed to find an effect of the cognitive load manipulation. These authors especially designed their experiment to test if the absence of cognitive load was due to the peripheral features needed for succeeding in the recall task not being refreshed. However, they found evidence against this hypothesis. They thus concluded that there were some boundary conditions to the cognitive load effect, although they could not pinpoint exactly what these conditions were.

Following the results from the three studies presented above, we can see that un-refreshable memoranda appear to pertain all in the visuospatial domain. However, it is difficult to find a link that ties the type of memoranda that do not seem to induce refreshing, aside of their belonging to the visuospatial domain. Two possible explanations for the existence of un-refreshable memoranda have been put forward: either (1) as proposed by Ricker and Cowan (2010), only memoranda that have a sufficient representation in long-term memory can be refreshed, or (2) only memoranda that have been sufficiently consolidated could be refreshed. For now, we will draw a temporary conclusion about attentional refreshing in working memory, and then we will investigate consolidation processes in working memory.

2.2.4 Conclusion on attentional refreshing

As analyzed above, attentional refreshing has been put forward to explain attentional effects in working memory. Two distinct conceptions of attentional refreshing presently exist in the literature, and it may be that both processes exist in parallel and support working memory maintenance. As said earlier, the main difference between both kinds of attentional refreshing relies on their deliberateness. Controlled refreshing is defined as conscious and active whereas swift refreshing is defined as fast and covert. Another way of differentiating both kinds of refreshing would be to describe controlled refreshing as a higher-level strategy to enhance maintenance of information in working memory, whereas swift refreshing is a semi-automatic mechanism that lies outside of conscious control. However, as we have indicated, attentional refreshing does not seem to be as general as previously thought. Some features in the visual domain seem to be impossible to refresh. Foreshadowing later parts of this thesis, we experimentally investigated the possible limits to swift refreshing in chapter 4.1 and 4.2, and controlled refreshing in chapter 6. In the next part of this chapter, we will describe consolidation as another attention-based process that influences recall performance in working memory.

2.3 Consolidation in working memory

In addition to attentional refreshing, consolidation has been described to explain several findings regarding attention after encoding in working memory.

Consolidation is a process in working memory that is thought to transform transient sensory traces (also called iconic memory traces) into stable working memory representations (De Schrijver & Barrouillet, 2017; Jolicoeur & Dell'Acqua, 1998; Ricker & Hardmann, 2017; see Ricker, Nieuwenstein, Bayliss & Barrouillet, 2018, for a review). This process is separated from encoding processes, as it relies on representations that are no longer present to our senses, and thus can be described as a working memory process.

First evidence for the existence of consolidation was based on the attention-blink paradigm. In this paradigm, two memoranda are presented in quick successions for later recall. The classical effect is that performance for the second memoranda will be worse if it is presented in a 200-500 millisecond window following the presentation of the first memoranda. This is called the attentional blink effect (Broadbent & Broadbent, 1987; Raymond, Shapiro & Arnell, 1992; Shapiro, 1997). These authors proposed that an attentionally demanding process was at work during this window, which would interfere with the processing of the second memoranda. However, the function of this process was still not evident at that point.

This led Jolicoeur and Dell'Acqua (1998) to investigate this proposed attention-based process using a paradigm closer to the working memory literature. In their seminal article, the authors used dual-task methodology to investigate this proposed construct. In seven experiments, participants had to memorize a visual display containing letters or symbols. Then, after a variable stimulus onset asynchrony (SOA), an auditive signal was sent to the participants. They had to respond to this signal as fast as possible. The authors manipulated the quantity of memoranda in the visual display and the SOA. The results showed that the reaction times to the auditive task depended on the memoranda's set size and the auditive signal SOA. Reaction times

to the concurrent task were increased during short SOA and decreased monotonically with larger SOA. Furthermore, this effect was more pronounced the higher the number of memoranda to be maintained. Interestingly, the effect of the SOA manipulation on the reaction times to the auditory signal was absent when participants were tasked to ignore the presented memoranda. Thus, the authors argued for the existence of an attentionally-demanding process that would postpone processing of the auditory signal. The authors supposed that this process would help the maintenance of information in working memory by transforming fleeting iconic memory traces into stable working memory traces. They also surmised that this process was a processing bottleneck, as reaction times to the auditory task would not have been postponed otherwise. Follow-up studies replicated the postponement induced by working memory consolidation and expanded our understanding of this mechanism. For example, Stevanovski and Jolicoeur (2007) showed that consolidation effect on reaction times was not dependent on the use of the phonological loop, which rules out the implication of verbal process in the effect.

Therefore, the first evidence of the existence of a consolidation process relied on the effect of SOA on the reaction times to a concurrent task. This showed that an attention-demanding process was at work during the short period following item presentation. However, these results did not show that this process actually enhanced working memory representations. Using the same dual-task methodology but with more difficult tasks, Nieuwenstein & Wyble (2014) showed an impact of the SOA manipulation on recall performance, with better performance linked to longer SOA, in addition to the classical effect on the reaction times to the concurrent task. This impact of the consolidation manipulation on recall performance was also found with other experimental paradigms closer to the working memory literature. For example, Ricker and Cowan (2014) investigated consolidation by using a series of masked simple span experiments. In this series of experiments, participants had to memorize a set of memoranda, presented either simultaneously or sequentially. Each memorandum presentation was always followed by a visual mask, which always

appeared 100 milliseconds after the memoranda disappeared from screen. Then, after a variable retention interval, a probe was shown on screen and participants had to decide whether it was previously presented or not. The results showed that having more time to consolidate a trace in working memory leads to a slower rate of temporal decay during the retention interval. Following these results, the authors argued that consolidation was happening during the retention interval. This process would make the working memory traces less susceptible to temporal decay and interferences, and possibly more easily attentionally refreshed. Thus, more time to consolidate a representation in working memory would lead to better recall performance, and possibly enhance the impact of attentional refreshing opportunity manipulation.

2.3.1 Separability between consolidation, encoding and effect of masked presentation

The important aspect of the Ricker and Cowan's study (2015) is its use of masked presentation. Since the consolidation manipulation relies on the presence of variable free-time just after memoranda presentation, it may be that encoding processes sustain the benefit attributed to this free-time. However, several findings argue against this possibility. First, it has been shown that consolidation is independent of visual characteristics of the memoranda (Sun, Zimmer & Fu, 2011). If encoding and consolidation were supported by the same processes, we should see an impact of perceptual features on consolidation, as it is evident for encoding processes. Secondly, recent studies showed that different aspects of the memoranda could be consolidated separately (Chen & Wyble, 2015, 2016). In their studies, Chen and Wyble asked participants to remember the locations of a letter in an array of numbers. Then, they asked the participants to recall the locations and the identity of the letters. Results showed that participants could easily recall the locations of the letters, but not their identity if they were not asked in advance to remember them. This was replicated in numerous experimental configurations, showing that this pattern of

results was not dependent on the material used. The pattern of results shows that identification, which requires encoding, does not always lead to consolidation. Thus, taken together, these results are congruent with a general separability of consolidation and encoding processes.

One way to distinguish the effect of encoding processes from consolidation would be to make use of masking procedure. Encoding relies on iconic memory (Tripathy & Öğmen, 2018) and masking overwrites the iconic memory trace of the just presented items. Therefore, using a mask after the presentation of each memorandum, and manipulating the SOA after the mask have been presented would insure that only the consolidation process would be manipulated. However, there is a possible confound with the use of masked presentation. Several studies found an impact of mask's SOAs on recall performance (Blalock, 2013; Vogel, Woodman, & Luck, 2006). It may be that masking a memorandum would stop its consolidation, and thus explain the relationship between masking and recall performance. If this is the case, then using masks would not be the solution to investigating working memory consolidation. Although the time course of the impact of masks on recall performance seems to indicate that consolidation is impacted by the masking procedure, several studies found an effect of consolidation delay manipulation, whether presentation was masked or not (e.g., Jolicoeur & Dell'Acqua, 1998; Nieuwenstein & Wyble, 2014; Ricker & Cowan, 2014).

Thus, consolidation probably interacts with masking presentation, but does not completely stop it. Two possible mechanisms have been described: either the masks interfere with the representation being consolidated (Ricker, 2015) or masking elongate the time needed for consolidation (Wyble, Bowman, & Nieuwenstein, 2009; Vogel, Woodman & Luck, 2006). However, irrespective of the kind of interactions between masking effect and consolidation, masking is useful in investigating consolidation, because it ensures that the effect of the free time post-mask presentation relies on working memory and not iconic memory.

2.3.2 Separability between consolidation and attentional refreshing

Both consolidation and attentional refreshing have been proposed as attention-based processes happening during maintenance in working memory. In addition, both processes use the same kind of attention. This could therefore imply that both processes are in fact one more general attentional process. Currently, only swift refreshing has been investigated orthogonally to consolidation. For this reason, we will focus the discussion on the relationship between swift refreshing and consolidation.

The study from Bayliss, Bogdanovs and Jarrold (2015) aimed at disentangling the possibility that attentional refreshing and consolidation were separate mechanisms or not. In Experiment 1 of their study, the authors created a complex-span task with letters as memoranda, in which they manipulated a delay in between item presentation and the beginning of the concurrent task (hereafter: consolidation delay) as well as the cognitive load of the concurrent task. They reasoned that, if consolidation and attentional refreshing were based on the same attentional process, then the cognitive load effect should be bigger in the condition with a longer consolidation delay, compared to a condition with a shorter consolidation delay. This would be marked by a statistical interaction between both independent variables. The results showed a significant effect of both variables, but found no evidence for the interaction. The results were replicated in following experiments. Experiment 2 found the same pattern of results with another concurrent task and Experiment 3 showed that the presence of an articulatory suppression method does not change the pattern of results. The authors thus concluded that attentional refreshing and consolidation are two separate mechanisms.

However, it is important to note that, in these experiments, the authors did not make use of a masked presentation paradigm. Therefore, their manipulation of consolidation delay may have been confounded with encoding processes. Nonetheless, this pattern of distinguishability between attentional refreshing and

consolidation, based on the absence of interaction between consolidation and attentional refreshing, was replicated by De Schrijver and Barrouillet (2017), but only when participants were under articulatory suppression. In their study, they also manipulated consolidation delay and attentional refreshing orthogonally in a complex-span task, but this time they had more than two levels of consolidation delay and a mask followed each memorandum presentation. In Experiment 1, both cognitive load and consolidation delay manipulations had an additive effect on recall performance, and there was evidence for the interaction. The effect of consolidation SOA manipulation progressively decrease as the cognitive load gets lower. However, De Schrijver and Barrouillet (2017) found that the presence or absence of articulatory suppression did change this pattern. In a control experiment, participants underwent the same procedure as experiment 1, but this time they were under articulatory suppression. Results from the control experiment were similar than experiment 1, with the main difference that evidence for the interaction disappeared. The authors concluded that the interaction between attentional refreshing and consolidation manipulation detected in experiment 1 was attributable to verbal rehearsal mechanism, and thus both attention-base processes were effectively separable. In addition, the authors of this second study also observed that, when the overall time for both consolidation and attentional refreshing processes was equated, recall performance was similar. Thus, these authors concluded that there is some substitutability between both processes.

Although both studies have slightly different conclusions, the fact that attentional refreshing and consolidation manipulation only interact when participants are not under articulatory suppression argue in favor of a separability between both mechanisms. This is also shown by the results from Bayliss et al. (2015) that free-time presented just after memoranda presentation is more beneficial than an equal quantity of free-time presented after the processing of a concurrent task. This advocates in favor of a special status of the free time happening just after item presentation. This special status is also evidenced by the attention-blink paradigm

presented at the beginning of this section and cannot be explained by the attentional refreshing mechanism. Thus, attentional refreshing and consolidation should be viewed as separable mechanisms.

Interestingly, attentional refreshing and consolidation may be associated in another way. It has been proposed that attentional refreshing can only be applied on representations that have been sufficiently consolidated (Morey & Cowan, 2018; Ricker & Vergauwe, 2020). We already noted regarding attentional refreshing (chapter 2.2.3) that some visuospatial material could not be refreshed (Ricker & Cowan, 2010; Ricker & Vergauwe 2020; Vergauwe, Camos & Barrouillet, 2014). Ricker and Vergauwe (2020) argued that one possibility to explain the absence of cognitive load and memory load effects with specific materials could be due to these items not being sufficiently well consolidated in working memory. Although we could say that results from Bayliss et al. (2015) and De Schrijver and Barrouillet (2017) argue against this possibility, two possible confounds could hinder this conclusion, and both rely on the fact that letters were used as memoranda, both in De Schrijver and Barrouillet (2017) and Bayliss et al. (2015).

First, as we saw in the chapter dedicated to attentional refreshing, only visuospatial materials were found impossible to refresh. This hints to the possibility that specificities about how visuospatial material is represented in working memory make their representations intrinsically more difficult to refresh. Secondly, it is much more common to have to decide the identity of a letter and maintain it for later recall than doing the same for fonts of the same letter (Vergauwe, Camos & Barrouillet, 2010) or for unconventional characters (Ricker & Cowan, 2010). Since there is a difference in familiarity between letters (which are highly familiar) and unconventional characters (which are, by definition, less familiar), it may be that consolidation relies, at least in part, on information held in long-term memory to function. Memoranda with better long-term representations could be consolidated more efficiently than

memoranda with weaker long-term memory representations. To foreshadow the experimental part of our thesis, this last possibility has been investigated in chapter 5.

2.3.3 Limits to consolidation

Consolidation has been less investigated than attentional refreshing in working memory settings. Thus, less is known on the limits between which consolidation process can happen. As we saw earlier, consolidation seems separated from attentional refreshing. It has been investigated by manipulating free time following item presentation, with the idea that the free time was used by consolidation to enhance the representation and somewhat shield it from decay and interference. The experimental paradigms used to investigate consolidation all rely on SOA manipulation between memoranda presentation and “something” that stops consolidation. In addition, a mask should follow each memorandum presentation to ascertain that consolidation is manipulated and not encoding. Indeed, masking should be used to be sure that free time is used for consolidation process, and not an encoding process, by replacing the content of iconic memory (and thus ensuring that working memory is at play during the free time). This means that masks should appear on screen after a fixed interval following memoranda presentation, irrespective of the consolidation delay condition and that free time should be manipulated after the mask disappeared from the screen.

Thus, the question raised is how to manipulate effectively a consolidation delay. One solution relies on the type of attention used by consolidation. We saw that consolidation uses the same attentional resource as attentional refreshing and other tasks demanding central attention. Thus, consolidation could be stopped if participant have to execute a concurrent attentionally demanding task. We could adapt complex-span tasks to manipulate the free time between mask presentation and the start of a concurrent task. This approach could enhance investigation of consolidation because we could apply the same logic as for attentional refreshing. Indeed, one could manipulate consolidation delay, memoranda material and

secondary tasks uses of central attention orthogonally. This could lead to a better understanding of consolidation and its possible limitations.

2.3.4 Conclusion about consolidation in working memory

Evidence for the existence of a consolidation process in working memory has been gathered during the last 20 years. Although the actual mechanism of functioning is still a matter of debate, its role has reached a consensus: it is a process that transforms iconic memory traces into a more stable working memory representation. Another aspect that has reached consensus is that this process needs central attention for its functioning. Since the process is effective after memoranda presentation, we could argue that it is comprised in the maintenance period of working memory. Although separated from attentional refreshing, both processes rely on the same resource, and thus function sequentially from one another, consolidation first, due to the processing bottleneck in working memory. This led researchers to suppose that attentional refreshing may work better with better-consolidated memoranda, compared to less-well consolidated memoranda. Also, consolidation, in the same vein as attentional refreshing, seems to be under conscious control, as manipulation of consolidation delay did not have an impact on concurrent task if participants were not tasked to remember it. Whereas the impact was evident when participants had to recall the presented memoranda. The same argument can be made regarding attentional refreshing, as memoranda are only maintained when participants are tasked to recall them. However, several aspects of consolidation functioning are still not known. For example, it is not yet clear if consolidation works in an all-or-nothing matters or if it is a more gradual process. Existing evidence points in the direction of a more gradual process, as De Schrijver and Barrouillet (2017) described a monotonic and continuous trend between recall performance and consolidation manipulation. More work is still needed to clarify this possibility. In the same vein, not much is known on the limitation to consolidation. It may be that consolidation is a very basic mechanism and that all memoranda that

enter working memory have to be consolidated. Or, perhaps, it is a more “peripheral” mechanism that enhances specific memoranda, depending on the context and task instruction. All in all, future research should aim at disentangling these different possibilities about consolidation, as its inner functioning could allow us to enhance our understanding of working memory.

2.4 Conclusion about attentional processes in working memory maintenance

Two separate attention-based processes are described in the literature to explain attentional effects during working memory maintenance: attentional refreshing and consolidation. Attentional refreshing is thought to enable domain-general maintenance of information by focusing central (or executive) attention on representations currently held in working memory. This boosts their activation level and thus protects them from forgetting via temporal decay and interferences. Consolidation is also described as an attention-demanding process, but its function is to transform fleeting iconic memory traces into stable working memory representations that can be more easily maintained in working memory. Although both processes rely on the same attentional resource, they both seem to work serially, one after the other. An item is firstly consolidated into working memory, then it is maintained via attentional refreshing.

Results from experiments manipulating both processes orthogonally tend to support an additive effect of attentional refreshing and consolidation. This is congruent with the idea that both processes are, at least in part, separated. Therefore, attentional refreshing will enhance the maintenance of information in working memory irrespective of the level at which it has been consolidated. It is not clear yet whether consolidation works in a more continuous manner or is more akin to an “all-or-nothing” process. For example, studies from Wyble and Chen (2016,2017) showed that specific aspects of the memoranda could be consolidated when other aspects were discarded. These results could be explained by an “all-or-nothing” consolidation process, where only the important features of the memoranda are consolidated over a threshold and thus are effectively maintained. In contrast, results from De Schrijver and Barrouillet (2017) are more in line with consolidation being deployed in a more continuous manner. More work is needed to disentangle these different possibilities about consolidation functioning.

Regarding attentional refreshing, several aspects of its functioning are still a matter of debate. Firstly, there is no consensus on the time scale at which refreshing occurs and the schedule used to choose which representation to refresh. Proponents of the controlled refreshing argue that attentional refreshing is a relatively long process (at around 500-600 milliseconds) that is essentially sequential and conscious in nature. Participants can actively choose on which representation refreshing will be deployed and thus the "schedule question" is trivial. For the proponents of swift refreshing however, attentional refreshing is short (at around 30-50 milliseconds) and lies outside of direct and conscious control. Thus, the schedule question is much less trivial. Investigating its time course could inform on the inner working of swift refreshing.

Secondly, it is not clear which kinds of memoranda cannot be refreshed. All memoranda that failed to induce a cognitive load effect were in the visuospatial domain. This hints to the possibility that some specificities about visuospatial memoranda render them impossible to be refreshed, when compared to verbal memoranda that were always altered by cognitive load and memory load manipulation. One possible explanation for this discrepancy between verbal and visuospatial materials could rely on long-term memory factors; perhaps only representations that are sufficiently well represented in working memory can be refreshed. Since letters and words are famously well represented in long-term memory, compared to specific peripheral visuospatial features, this could explain why only visuospatial memoranda failed to be refreshed.

The next chapter of our thesis will investigate the relationships between long-term memory and working memory from an attentional point of view. We will first evaluate more closely the effects of long-term memory factors on working memory functioning. Then, we will present methodological problems to manipulate semantic long-term memory factors in working memory. Finally, we will present recent studies

that investigated the possibility that long-term memory directly influences attentional refreshing and consolidation.

Chapter 3: the relationships between working memory and long-term memory from an attentional point of view

In previous chapters, we focused our investigation on working memory and attention. But, as noted in our introduction, one of the main debates in the working memory literature is the nature of its relationship with long-term memory. We saw that some models consider working memory as a separate store from long-term memory, whereas some other models consider working memory as the activated part of long-term memory. However, all modern models acknowledge that both systems must interact heavily. The debate about the relationships between both systems could fuel several theses on its own. Thus, in this chapter, we will focus on the impact of long-term memory factors on working memory functioning, with a focus on attentional processes during maintenance. We will first describe general effects of long-term memory on working memory functioning, then we will explore more precisely how to manipulate long-term memory in a working memory setting. Finally, we will investigate the possibility that attentional refreshing and consolidation, presented in the previous chapter, might be influenced by long-term memory factors.

3.1 Effect of long-term memory factors in working memory

Before investigating the effects of long-term memory factors in working memory, one has to define what is long-term memory. Often described in opposition to working memory, which is a time limited and capacity limited system, long-term memory is a seemingly capacity-free system that can indefinitely hold information in a passive format (Cowan, 2008). This information can be transferred or reactivated (depending on the theoretical background) into working memory to assist ongoing cognitive or behavioral activity. This system is not unified. There is consensus on distinguishing between explicit memory and implicit memory (also called declarative and procedural memory, respectively, Aggleton, 2008; Meulemans & Van Der Linden, 2003). Explicit memory is the system that holds information that can be consciously

and verbally described, thus the term “declarative” memory. For example, remembering what the person ate two days ago relies on explicit memory. In contrast, implicit memory comprised mostly of memory for movements of the body and how to use objects, which relies mostly outside of direct conscious control and seems to be based on different brain structures than declarative memory. How to tie shoes is an example of a memory held in procedural memory.

In this thesis, we will focus on declarative memory. More precisely, declarative memory can itself be decomposed in at least two subdivisions. On one hand, there is episodic memory. It is comprised memory of things that actually happened to the person (Ranganath, Michael & Craig, 2005). For example, remembering what was done last weekend or the person’s birthday party ten years ago relies on episodic long-term memory. On the other hand, there is semantic memory, which refers to decontextualized factual knowledge about the person’s world (Wood, Baxter & Belpaeme, 2011). For example, knowing that the capital of France is Paris, which side is north on a map, or the meaning of a word are all information held in semantic long-term memory. In our thesis, we focused our investigation on the effect of semantic long-term memory on working memory functioning.

Although working memory was first described as a gateway for entering long-term memory (e.g., Atkinson & Shiffrin, 1968), the inverse relationship (effect of long-term memory factors on working memory functioning) became evident shortly afterwards. One of the first pieces of experimental evidence for the impact of long-term memory on working memory functioning came from studies investigating the chunking effects (Chase & Simon, 1973; Miller, 1956). In the seminal study from Chase and Simon (1973), participants had to memorize the location of chess pieces coming from a legal middle-game position (chess pieces in positions that could “naturally” emerge from real games following chess rules). The authors then compared recall performance between beginner chess players and advanced or master level players. The beginner players could recall the location of around four pieces, which

corresponds to the classical capacity of working memory. In stark contrast, advanced and master level players could recall around sixteen pieces. Since all memorized positions came from real games of chess, the authors concluded that advanced and master chess players could use their knowledge about chess and the movements of chess pieces, held in semantic long-term memory, to enhance working memory capacity. The authors argued that instead of memorizing the location of each piece separately, the advanced and master player memorized patterns of locations that are often found during chess game. Each of these patterns would be maintained in working memory as a unified chunk comprising several pieces each. This explanation has been evidenced by results from Frey and Adelman (1976), which showed that the advantage for advanced and master chess players to memorize locations of pieces was dependent on the legality of the piece locations. When the chess pieces were in patterns that could not be possible following chess rules, the recall advantage for advanced and master players was greatly reduced. This means that advanced and master chess players could actively use their knowledge of chess to chunk several pieces together and thus memorize a higher total number of pieces than beginners, who could only maintain chunks containing a single piece each.

Taken together, these results show one of the first experimental exemplifications of the chunking mechanism: several pieces of information can be fused together to become one representation, and thus enhance working memory capacity by memorizing chunks of information containing more pieces instead of each piece separately. Said differently, chunking is the fact of recoding several smaller units of information into larger units and then only remembering these larger units. This mechanism seems to be domain general, as it affects visuospatial as well as verbal material (Cowan, Rouders, Blume & Saults, 2012).

The chunking effect shows that information held in semantic long-term memory can directly affect working memory functioning. Moreover, several other similar effects of semantic long-term memory factors on working memory functioning have

been described in the literature. For example, lists of words that are semantically close to each other (e.g., leaf-tree-wood) will be better recalled than list of words which are not semantically related (e.g., house-sky-cork). This is called the semantic relatedness effect and have been reliably replicated (Poirier & Saint-Aubin, 1995; Saint-Aubin, Quелlette & Poirier, 2005). This shows that information in long-term memory is organized in a semantic manner: representations held in working memory will activate to some level representations that are semantically linked to it. These semantically related representations would thus be more accessible than non-semantically related representations. If we take the embedded-processes model of working memory approach, we could say that representations held in the focus of attention will automatically activate semantically related representations in the region of activated long-term memory. This is also evidenced by the semantic priming effect, in which words semantically related to a probed word will be responded to more quickly after probe presentation than word not semantically related to the probe (Monsell, Doyle & Haggard, 1989; Sperber, McCauley, Ragain, & Weil, 1979; Voyer, 2003; see Weingarten, Chen, McAdams, Yi, Hepler & Albarracín, 2016, for a review).

Two other examples of the impact of long-term memory factors on working memory functioning are the lexical frequency and the lexicality effect. The lexical frequency of a word is a measure of how common a word is in a specific language. It is usually computed by counting the number of occurrences of a word in a corpus of text. For example, the French words database "Lexique3" (New, Palier, Brisbaert & Ferrand, 2004) gives two numbers for the lexical frequency of French words, one based on a corpus of books whereas the second is based on a corpus of subtitles from movies. The number given is in occurrence per million of words. The lexical frequency effect in itself is the fact that words with higher lexical frequency words are processed more quickly and induce better recall performance than words with lower lexical frequency (Gardner, Rothkopf, Lapan, & Lafferty, 1987; Rubenstein, Garfield & Amilukan, 1970; Scarborough, Cortese & Scarborough, 1977). This effect is closely related to the lexicality effect: For native expert speakers of the target language, real

words from a target language are more precisely and more quickly responded to when compared with pseudo-words (words that follow the phonotactical rules of the target language but have no proper meaning; Hulme, Maughan & Brown, 1991; Hulme et al. 1997; Majerus & Van der Linden, 2003). Indeed, pseudo-words can be thought of as words with a lexical frequency of zero. In the following part of this chapter, we will use the term “lexical frequency effect” to refer to both the lexical frequency and the lexicality effect.

The effects presented above induced two main conclusions. Firstly, representations in long-term memory differ on the strength of their encoding. Information in long-term memory does not seem to have an “all-or-nothing” encoding status, but a more continuous representation level. Secondly, information that is better represented in long-term memory induces better recall performance and is recalled more quickly than information that is less well represented in long-term memory. Indeed, all the effects presented above point in the same direction: the more a memorandum is encountered in day-to-day lives, the better it will be encoded in long-term memory. This logic also explains the difference between beginners and advanced chess players on the number of pieces recalled: advanced players have had much more experience with chess boards, even from a perceptual point of view, and thus can use their more deeply ingrained chess knowledge to chunk information together.

One important aspect of the effects presented above is their apparent automaticity. It could be argued that lexical frequency effect and chunking effects could come from the strategical use of long-term memory. However, results from recent studies contradict this conclusion. For example, Kowialiewsky and Majerus (2018) showed that, when minimizing the opportunity to engage in strategic processes, the semantic relatedness and the lexical frequency effects were still evident. Thus, it seems that at least these two effects come directly from the architecture of long-term memory and its link to working memory.

To conclude, the effects described above show that long-term memory factors have a direct impact on working memory functioning. Although these effects have been reliably replicated; it is not actually known by which mechanism these effects arise. The exact locus of these effects in working memory is still debated. It could be that encoding processes are more efficient for information better represented in long-term memory, or it may be that less well-encoded information in long-term memory induces slower retrieval during the recall phase, or many other possibilities. In the next part of this chapter, we will investigate how to actually manipulate long-term memory in a working memory setting. We then will investigate the hypothesis that attentional refreshing and consolidation processes during maintenance are at the basis of the effect of long-term memory factors on working memory functioning.

3.2 How to manipulate long-term memory in a working memory setting

Before investigating attentional processes during maintenance and their link to long-term memory, we will examine the manipulation of long-term memory factors in working memory setting. We showed above (chapter 3.1) that early on researchers aimed at manipulating aspects of long-term memory into the working memory paradigm (e.g., Chase & Simon, 1973; Miller, 1956). The general idea was that the strength at which information is encoded in long-term memory was dependent on the amount of time the memoranda were encountered in day-to-day life. However, this leads to another problem of evaluating at which degree some items are more present in day-to-day life compared to another items.

Two main solutions have been described in the literature to resolve this methodological problem. One solution is to compare a population known in advance to have encountered a specific class of memoranda more often than another population, and then compare their respective performance on working memory tasks involving this specific class of memoranda. This is exactly the approach used by Chase and Simon (1973) when comparing expert and beginner chess players. This solution has the advantage to be more ecological, because the memoranda used are usually taken from real life situation (e.g., birds, chess game, well-known acronyms, cars). However, this method also has to rely on a non-random between-subject design. Although each participant can have more experimental trials for the same experiment duration than a within-subject design (because participants do not do trials in all the experimental cells), other variables intrinsic to both populations could also be confounded. For example, comparing a population of medical doctors and retail workers on their ability to maintain specific visual patterns often found in radiography, not only entails measuring a difference in the groups differing expertise on this specific subject, but also possible differences linked to other aspects of the groups. For example, retail workers will likely have less years of higher education than medical doctors. Indeed, it has been shown that people that spend more years in

higher education tend to better results in working memory capacity tests (Souza-Talarico, Caramelli, Nitrini, & Chaves, 2007). Thus, if differences in working memory performance between both groups are detected, at least some part of this difference could be due to individual differences in working memory capacity, which in turn can be impacted by socio-economic aspects of the participants. In general, this is a limitation of all experiments that use non-random groups and between-subject design. Due to the complexity of human cognition, several confounds will almost always be present when comparing two groups of ethnically, socio-economically or culturally different people in working memory or long-term memory tests.

The second solution to manipulate the long-term memory factor in a working memory setting is to present, to the same group of participants, memoranda which intrinsically vary on their propensity to be well-encoded in long-term memory. This solution avoids the risk induced by non-random between subject design described above. However, it means that this propensity to be well encoded in long-term memory has to be known in advance and evaluated separately from the experimental setting. In the verbal domain, this solution has been reliably used to index long-term memory in a working memory setting by using the lexical frequency of words to recall. As noted above, words with higher lexical frequency induce better recall performance and are responded to more rapidly than words that have a lower lexical frequency. Since lexical frequency is a proxy measure of how much some words are encountered in comparison to other words in day-to-day life, it means that effect induced by the lexical frequency manipulation relies on the depth of encoding of these words in long-term memory. This solution seems more elegant than the non-random between-subject manipulation presented above, because much more intra-personal and inter-personal factors influencing performance can be neutralized or controlled in an experimental setting.

However, it raises the issue of advance knowledge on which kind of memoranda are encountered more often in day-to-day life compared to other. Thanks to the work

of lexicology, several databases of words and their lexical frequency (in addition to many other linguistic factors) have been computed for huge corpus of words (e.g., Lexique3, New et al., 2003, for French words). These databases are helpful to create list of words that differ on lexical frequency but are otherwise equivalent on other linguistic factors (number of syllables, consonant/vowel structure, number of phonological neighbors, etc.) and thus permit the evaluation of subtle aspects of the memoranda. Furthermore, since lexical frequency is a number comprised between 0 and 1'000'000, it is possible to manipulate lexical frequency (and thus long-term memory) in a more continuous manner. In contrast, it is much more difficult to manipulate long-term memory in the visuospatial domain via the within-subject solution.

To our knowledge, there has been no real attempt to create a direct equivalent of lexical frequency in the visuospatial domain. The problem is that lexical frequency can be considered as an intrinsic feature of each word and can be computed outside of experimental manipulation. On the contrary, in the visuospatial domain there is no corpus of images that can be directly used to measure a similar index.

In this domain, researchers wanting to manipulate long-term memory tended to use an effect specific to the visuospatial domain: the own-race effect (ORE). The ORE is the fact that people from a specific ethnicity will recall and respond to faces of the same ethnicity more efficiently than faces from another ethnicity (Levine, 1996; Valentine & Endo, 1992; See Johnston & Edmond, 2009, for a review). For example, Swiss-based Caucasian people will remember the faces of Swiss-based Caucasian people better than Japanese faces, while Japanese people will remember Japanese faces better than Swiss Caucasian faces. It has been shown that the difference induced by the ORE manipulation depends at least in part on the fact of seeing faces from the own ethnicity much more often than faces of other ethnicities. For this reason, the ORE can be considered as a kind of perceptual expertise (Tanaka, Heptonstall & Hagen, 2013). The idea is that regularities about faces from the own

ethnicity will be better encoded in long-term memory than specificities of faces from other ethnicities. Since this is the definition of long-term memory manipulation, several studies used this effect as a basis to manipulate long-term memory in the visuospatial domain (Stahl, Wiese & Schweinberger, 2008; Stelter & Degner, 2018).

Another way to manipulate long-term memory via the use of faces is to present two groups of faces, one comprising faces of random unknown people, whereas the second group comprised of faces of well-known people (actor, musician, television host, etc.). This manipulation between familiar faces and unfamiliar faces has established that familiar faces are recognized more easily and faster than unfamiliar faces (Burton, Bruce & Hancock, 1999; Hancock, Bruce & Burton, 2000; see Natu & O'Toole, 2011, for a review). Like the ORE, this effect is dependent on the frequency at which a face has been encountered prior to the experiment, and thus depends on a difference in long-term memory representation.

One last way to manipulate long-term memory would be to train participants on specific memoranda, and then compare their performance with participants that were not trained on the memoranda. If participants are attributed randomly to one of the group, this would limit the possible confound induced by non-random between subject design, while ensuring that both groups differ on their representation of the memoranda in long-term memory. This solution would be useful in the visuospatial domain, as there is no existing corpus of memoranda that we know in advance differ on their long-term memory representation.

To conclude, there are several ways to manipulate long-term memory in a working memory setting. However, depending on the actual research question and the domain of investigation, several methodological problems can arise. In the verbal domain, the most straight forward long-term memory manipulation relies on the manipulation of lexical frequency which can be computed separately and a-priori from the experimental setting. The visuospatial domain, however, lacks such clearly-defined features. To investigate long-term memory in this domain, researchers relied

mostly on effects regarding face processing. However, it seems that faces have a special status in the cognitive system and are processed differently than other visuospatial memoranda (Crist, Wu, Karp & Woldorff, 2008; Yin, 1969). Thus, conclusions relying on effect using faces cannot be directly generalized to the visuospatial domain in general. To foreshadow the experimental part of this thesis, we investigated how to manipulate long-term memory in the visuospatial domain in chapters 4 and 5.

3.3 Attentional refreshing, consolidation, and long-term memory

As indicated above, attentional refreshing and consolidation in working memory have been proposed to explain attentional effects during working memory maintenance. Although much is to be discovered on the operation of both processes, it may be possible that they rely, at least in part, on long-term memory representations. The next sections of this chapter will be dedicated to investigating this possibility, separately for attentional refreshing and consolidation.

3.3.1 Attentional refreshing and long-term memory

Proponents of the TBRS model of working memory have supposed that attentional refreshing may use atomic elements from long-term memory to reconstruct partly disintegrated traces (Barrouillet & Camos, 2015). In their view, refreshing could work similarly to a redintegration process. Proposed by Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart (1997), a redintegration process is a process via which representations in working memory are reconstructed using information held in long-term memory. For example, when hearing a few notes from a favorite song and then recalling the whole song from long-term memory is a redintegration process. In regard to attentional refreshing, this would mean that elements of information held in long-term memory would be used to reconstruct the memory traces in working memory which have been partly deactivated due to temporal decay and interferences. The information used for this process would be information that is semantically linked to the traces being reconstructed.

If attentional refreshing directly uses long-term memory information for reconstructing refreshed working memory traces, one could suppose that information better represented in long-term memory would be refreshed more efficiently than information not represented as well in long-term memory. Camos and colleagues (2019) directly investigated this hypothesis. In a series of four experiments, the authors manipulated lexical frequency and lexicality in a complex-span task orthogonally to cognitive load in Experiment 1 and 2, respectively. Then, they

manipulated again lexical frequency and lexicality orthogonally to the number of memoranda to maintain (i.e., the memory load) in two Brown-Peterson tasks. They surmised that if attentional refreshing is more efficient for words that are better encoded in long-term memory (i.e., high lexical frequency words in Experiment 1 and 3; real words in Experiment 2 and 4) compared to words that are less well encoded in long term memory (low frequency words in Experiment 1 and 3; pseudo-words in Experiment 2 and 4), then the cognitive load manipulation should have a bigger impact on recall performance to well-encoded words compared to less well-encoded words in the complex-span tasks. Regarding the Brown-Peterson tasks, their hypothesis was that the memory load effect would be less marked for familiar words, compared to less familiar ones, because more familiar words would be refreshed more quickly. The expected pattern of results in both experiments would be marked by the presence of a statistical interaction between both independent variables (cognitive load manipulation and the lexical frequency/lexicality manipulation in the complex-span tasks; memory load manipulation and lexical frequency/lexicality manipulation in the Brown-Peterson tasks). The results contradicted the hypothesis: the cognitive load manipulation and the lexicality/lexical frequency manipulations only had an additive effect on recall performance in the complex-span tasks, while only an effect of memory load was present on the reaction times in the Brown-Peterson tasks. Thanks to the use of Bayesian statistics, evidence was gathered against the statistical interaction in all experiments. This means that the manipulation of semantic long-term memory did not influence the degree at which words were refreshed.

However, the conclusion of the study presented above is limited to the verbal domain. As indicated in the chapter dedicated to attentional refreshing (chapter 2.2), it has been shown that some visuospatial memoranda were not affected by cognitive load manipulation, whereas the effect of the cognitive load manipulation was always present in the verbal domain. We know that the reading capacity depends on a long learning process that takes years to master. Thus, it could be that, even though words

vary intrinsically on long-term memory aspects, they still are, on a basic level, pretty much well-encoded in long-term memory. Manipulating lexical frequency or lexicality may thus not be a strong enough manipulation to really create a difference on attentional refreshing deployment. To generalize the conclusion of independence between attentional refreshing and semantic long-term memory factors, similar constructs need to be tested in the visuospatial domain.

3.3.2 Consolidation and long-term memory

Consolidation in working memory has also been proposed to be influenced by semantic long-term memory factors. The consolidation speed hypothesis stipulates that more familiar memoranda will be consolidated more quickly than less familiar memoranda. At least two studies addressed this question. Blalock (2015) had participants learn some shapes, and then compared the performance of participants in a four alternative forced choice (4-AFC) change detection tasks in which they also manipulated the stimulus onset asynchrony (SOA) between memoranda presentation and a mask, and presented either the trained shapes or shapes that were not trained. The idea of the manipulation was that manipulating the SOA between memoranda presentation and the mask would correspond to the consolidation manipulation. They found that trained shapes (i.e., more familiar memoranda) were responded to faster, induced better recall performance, and were less affected by the masking manipulation than untrained shapes (i.e., unfamiliar memoranda). The authors thus concluded that memoranda familiarity had an impact on consolidation, where more familiar memoranda were consolidated more efficiently and quickly than less familiar memoranda.

Xie and Zhang (2015) addressed the same question, but they used a slightly different approach. To manipulate long-term memory, they used a more ecological approach via using images coming from real-life examples. They contrasted the familiarity of small drawn monsters coming from a very popular video game called "Pokémon". Their manipulation of familiarity was twofold: First, they contrasted

images of Pokémon coming from the first generation to images of Pokémon coming from the last generation at that time². Secondly, participants were separated in two groups differing on their overall familiarity with Pokémon: participants who were familiar with Pokémon (i.e., the familiar group) and participants who weren't (i.e., the unfamiliar group). Thus, they manipulated familiarity via the generation from which the memoranda came in addition to the group knowledge manipulation, where the two groups of participants differed on their basic knowledge of Pokémon. The authors used a two-alternatives forced choice (2-AFC) recognition paradigm, where four memoranda were presented simultaneously on screen, then, after a variable SOA, a mask appeared on screen. Finally, a probe appeared in one of the locations where the memoranda were presented, and participants had to indicate if the probe was presented at the beginning of the trial or not. The results showed that participants that already had knowledge of Pokémon (the familiar group) consolidated images of Pokémon coming from the first generation faster than Pokémon coming from the latest generation. In contrast, the Pokémon generation's manipulation had no impact on recall performance for participants that had little to no prior knowledge on Pokémon (the un-familiar group). Thus, the conclusion was the same as Blalock's (2015) study: consolidation was faster for more familiar memoranda compared to less familiar ones.

However, both experiments presented above had limitations regarding their conclusions. The main one concerns the manipulation they used to investigate consolidation. In both studies, consolidation is manipulated by varying the SOA between memoranda presentation and the appearance of a mask on screen. Yet, we indicated earlier that manipulating consolidation by varying the SOA in-between a

² Pokémon is a popular video game that first came out in 1998, and in which you capture and train small creatures, called Pokémon. It is very popular with youngsters. Several generations of Pokémon came out since then, one approximately every two years, which all introduced new Pokémon. The authors thus manipulated familiarity by contrasting Pokémon introduced in the first generation (with whom participants were familiar) and Pokémon introduced in the last generation, with whom they were not familiar.

memoranda presentation and a mask makes it difficult to disentangle between consolidation and the encoding process. Thus, the results presented above could be due to familiar memoranda being encoded more efficiently than the less familiar, and not relying on an actual difference in consolidation. Furthermore, the study Xie and Zhang's study (2015) also relies on a non-random between-subject design, which is subject to the limitations explained above (chapter 3.2).

To conclude, although studies investigating long-term memory and consolidation seem to conclude that memoranda better represented in long-term memory tend to be consolidated faster, several confounds need to be eliminated before a firm conclusion can be drawn.

3.4 Conclusion on the relationship between attention, long-term memory and working memory maintenance

The manipulation of long-term memory in a working memory setting and the current state of the art on the links between attentional refreshing, consolidation, and semantic long-term memory allow several conclusions. Regarding attentional refreshing, present evidence is more congruent with a functional separation from semantic long-term memory. However, there is currently no consensus on the subject. One of the main unknown aspects of attentional refreshing is the fact that some materials seem to be immune to the effect of the cognitive load manipulation. Since the cognitive load manipulation is the main evidence for the existence of swift refreshing, understanding what makes some material immune to cognitive load manipulations could be very informative on the inner functioning of attentional refreshing. For the moment, more research is needed to explore the possible limitations to attentional refreshing and generalize the present results.

In regard of consolidation, less is known about this supposed process compared to attentional refreshing. Present evidence is congruent with the idea that it is a domain-general process that uses executive (or central) attention. It may be that consolidation is more efficient for more familiar memoranda, although other confounds need to be ruled-out before confirming these results.

After these theoretical aspects, we will introduce the actual experiments that we ran to investigate more directly the links between semantic long-term memory and attention-based processes during working memory maintenance. In chapter 4.1 and 4.2, we investigated the hypothesis that familiarity in long-term memory will influence swift attentional refreshing efficiency in the visuospatial domain. In these two chapters, we used complex-span task and Brown-Peterson tasks in which we manipulated visuospatial familiarity orthogonally to cognitive load. In chapter 5, we asked the same question but regarding consolidation instead of swift attentional refreshing. We used complex-span in which we manipulated the delay between a

masked memoranda presentation and the beginning of the concurrent task. Finally, in chapter 6, we investigated controlled refreshing via the guided refreshing paradigm developed by Souza and colleagues (2014) and its possible link to long-term memory.

Experimental Part

Chapter 4.1: Visuospatial Familiarity Effect on Attentional Working Memory Maintenance

Working memory is defined as the part of our cognitive system dedicated to the maintenance of a limited amount of information for short duration as well as online processing of currently relevant information (Barrouillet & Camos, 2015; Cowan, 2017, Oberauer, 2009). It is one of the most extensively investigated topics in psychological science, as it is central to day-to-day functioning and has been linked to several higher cognitive functions (Conway, Cowan, Bunting, Theriault & Minkoff, 2002; see Conway, Jarrold, Kane, Miyake, & Towse, 2007, for a review).

One central debate about working memory is its relationship to semantic long-term memory, where a seemingly infinite quantity of semantic information can be stored for long-time storage and retrieval (Baird, Baird & Wittinger, 1975; Brady, Konkle, Alvarez & Oliva, 2008). On the one hand, some authors argue that semantic long-term memory and working memory are not functionally distinct, but that working memory is basically activated long-term memory (Oberauer, 2012; Cowan, 2008). On the other hand, other authors consider working memory as a system separate from long-term memory (Baddeley & Hitch, 1974; Baddeley, 2000; Barrouillet & Camos, 2015).

One possible avenue to investigate the link between working memory and long-term memory would be to explore how semantic long-term memory factors impact working memory maintenance processes. Accordingly, the present study focuses on one maintenance mechanism in particular, called attentional refreshing, which is an attention-based mechanism that maintains mnemonic traces active in working memory against decay and interference (see Camos, Johnson, Loaiza, Portrat, Souza & Vergauwe, 2018, for a review). Among the diverse conceptions of refreshing, one suggests that it relies (at least in part) on semantic long-term memory representations (Barrouillet & Camos, 2015). This conception supposes that attentional refreshing uses information stored in semantic long-term memory to

reconstruct working memory traces. Accordingly, one would expect that more familiar items would benefit from better reconstruction, in other words being more efficiently refreshed, than less familiar ones. Previous examinations of the influence of item familiarity on attentional refreshing efficiency used verbal materials and provided evidence against this hypothesis. However, this finding can be specific to verbal working memory. The aim of the present study was to re-examine this hypothesis by testing the effect of item familiarity on attentional refreshing in visuospatial working memory.

4.1.1. Attentional refreshing in working memory

Attentional refreshing is conceived as a domain-general maintenance process that uses central or executive attention to keep representations active in working memory for subsequent processing (Johnson, 1992; Camos et al., 2018). It is thought to work by increasing the activation level of just-presented information, thus prolonging its accessibility in working memory. This increase in activation is achieved by focusing attention on the just-presented information. To put it in another way, it is the act of briefly thinking of just-presented information, when this information is not directly available to our senses anymore.

At least two indexes have been used to investigate attentional refreshing. A first index relies on the effect of a concurrent attentionally demanding task on recall performance. In a complex-span paradigm where a memory task is interspersed with a concurrent processing task, it has been shown that recall performance is influenced by the attentional demand of the concurrent processing task. The more this concurrent task uses attention per unit of time, the worse is recall performance. This is called the cognitive load effect (e.g., Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; Barrouillet & Camos, 2012, for a review). According to these authors, the cognitive load effect results from the functioning of attentional refreshing, because a more attention-demanding concurrent task would divert more central

attention from refreshing the to-be recalled items, compared to a less attention-demanding concurrent task.

A second index used to examine the existence of attentional refreshing relies on an effect of memory load on response times to a concurrent processing task. Several studies showed that the mean response time to a concurrent task increases for each additional item to maintain in working memory. For example, Vergauwe, Camos and Barrouillet (2014) developed a paradigm inspired by the Brown-Peterson task (Brown, 1958, Peterson & Peterson, 1959). This task was similar to a complex span task, with the difference that item presentation was not interspersed by a concurrent processing task. Instead, the processing task lasted for a fixed duration between item presentation and recall. The authors reported that the response times to the concurrent tasks increased linearly with the number of memory items (i.e., the memory load; see Camos, Mora, Oftinger, Mariz-Elsig, Schneider & Vergauwe, 2019, for similar findings). Jarrold, Tam, Baddeley, and Harvey (2011) showed a similar pattern of results using a complex-span task where they analysed the response times to the concurrent task following each image presentation. In all these studies, the response times to the concurrent task were approximately 40ms slower for each additional item held in working memory. The effect of memory load on a concurrent processing task is considered as evidence for attentional refreshing. Since refreshing and processing cannot be done simultaneously, the processing of the concurrent task is postponed until all the items held in working memory are refreshed. The 40ms slope found in all studies is then interpreted as the time it takes our cognitive system to refresh one memory item.

Despite the extensive research on attentional refreshing, its inner working is still up to debate. Researchers agree that central attention is needed for refreshing (e.g., Vergauwe, Dewaele, Langerock & Barrouillet, 2012), but otherwise several different proposals about its functioning have been put forward. On the one hand, its functioning has been described as a mechanism that uses semantic long-term

memory representations to reconstruct working memory representation (akin to a redintegration process; Thorn, Gathercole & Frankish, 2005; Barrouillet & Camos, 2015). On the other hand, Vergauwe and Cowan (2015) described its functioning as a rapid scanning of the current content of working memory. The main difference between the two proposals lies in the involvement of semantic long-term memory knowledge. Thus, investigating to what extent semantic long-term memory factors can influence attentional refreshing could give us information on the inner functioning of attentional refreshing and, as such, on the relationships between semantic long-term memory and working memory more broadly.

4.1.2 The role of semantic long-term memory in attentional refreshing

As said earlier, some proposals on the functioning of attentional refreshing hypothesise that the mechanism relies on semantic long-term memory representations to support the maintenance of information in working memory. One way to test this hypothesis is to examine how attentional refreshing is affected by well-replicated effects showing an effect of semantic long-term memory factors on working memory functioning. Two effects were first examined in relation to attentional refreshing (Camos et al., 2019): the lexicality effect and the lexical frequency effect in verbal working memory. The lexicality effect is the observation that words are better recalled than non-words. The lexical frequency effect is the observation that more frequent words are better recalled than less frequent words. Although these effects have been replicated several times in working memory tasks (e.g., Hulme, Maughan & Brown, 1991; Majerus & Van der Linden, 2003), its locus in working memory is still up to debate.

One possibility is that these effects reflect a differential impact of attentional refreshing that depends on item familiarity, such that more frequent words are refreshed more efficiently than less frequent words, because their representation in long-term memory is better defined or more easily accessible. To test this hypothesis, a recent study manipulated the cognitive load orthogonally to either the lexical

frequency of words or the lexicality of memory items, in a complex span task (Camos et al., 2019). In a first experiment, series of low-frequency or high-frequency words were presented in a complex-span task in which the cognitive load of the concurrent task was also manipulated. The authors' rationale was as follows: if more frequent words are refreshed more efficiently than less frequent words, then the lexicity effect (the difference in recall performance between the low- and high- frequency words) should be larger in trials where participants had more time to refresh (low cognitive load trials) than in trials where they had less time to refresh (high cognitive load trials). This should result in a statistical interaction between the lexical frequency of memory words and the cognitive load of the concurrent task. Results showed evidence against this interaction of interest, contradicting the aforementioned hypothesis. A similar absence of interaction with cognitive load was observed in a second experiment in which the lexicality of the memoranda (words vs. non-words) was manipulated.

In two additional experiments, the authors tested the same hypothesis but used another index of attentional refreshing, i.e., the effect of memory load on response times to a concurrent processing task. In particular, Experiment 3 introduced low- and high-frequency words in a Brown-Peterson task, following the same logic as the task developed by Vergauwe et al. (2014). The idea was that if the memory load effect on the response times of the concurrent task indexes attentional refreshing and if high frequency words are better refreshed than low-frequency words, then an interaction should be detected between both variables. High-frequency words would postpone the concurrent task to a lesser extent compared to low-frequency words, because high-frequency words would be refreshed more quickly. The results contradicted this prediction, showing evidence against the expected interaction. This pattern of results was replicated in a final experiment, in which the lexicality of memory items was manipulated instead of the lexical frequency in the same Brown-Peterson task.

Together, these four experiments provided evidence against an interaction between semantic long-term memory factors and attentional refreshing, which casts doubt on the hypothesis of an involvement of semantic long-term memory in the functioning of attentional refreshing. However, all the experiments described above were limited to the verbal domain. Even though attentional refreshing has been described as a domain-general mechanism (Barrouillet & Camos, 2015; Cowan, 1995; Oberauer & Hein, 2012) and should thus operate similarly for verbal and visuospatial domains, this theoretical position has not reached a consensus.

4.1.3 Differences between verbal and visuospatial domains

The most influential model of working memory supposes a functional separation between verbal and visuospatial working memory (Baddeley, 1986, 2000; Baddeley & Logie, 1999). It could then be that, even though attentional refreshing is deployed in the same way for verbal and visuospatial information, specificities about how information is actually represented could differ between domains and impact the functioning of attentional refreshing. This possibility is supported by the findings of Ricker and Cowan (2010), which showed that some item features in the visuospatial domain are inevitably forgotten after a retention interval compared to words or letters that are less susceptible to temporal decay. The authors interpreted this pattern of results as possible evidence for a differential impact of attentional refreshing between domains, and that this discrepancy would be due to differences on how information is represented between the verbal and the visuospatial domain.

In addition, several other studies reported an absence of cognitive load and memory load effects on different visuospatial memoranda. In one experiment, Vergauwe et al. (2014) found no effect of memory load on the response times to a concurrent task in a Brown-Peterson paradigm where participants had to maintain distinct fonts of the same letter. In the same vein, Ricker and Vergauwe (2020) consistently failed to find an effect of cognitive load manipulation in Brown-Peterson in which participants had to maintain the location of a dot on the edge of a circle. In

all their experiments, the authors failed to find an effect of an attentional refreshing manipulation on specific visuospatial items and argued that this discrepancy with the literature could be due to the fact that the visuospatial memoranda used in these experiments lack stable semantic long-term memory representation, rendering them impossible to refresh. These findings imply that the visuospatial domain differs from the verbal domain, in which the cognitive and memory load effects have been consistently observed whatever the nature of the verbal memory items (see Barrouillet & Camos, 2015, for review). Moreover, the authors' suggestion to explain the lack of cognitive and memory load effects on these visuospatial memoranda implicitly acknowledges that the functioning of attentional refreshing relies, at least in part, on semantic long-term memory. This is congruent with Barrouillet and Camos (2015)'s proposal on how attentional refreshing uses information stored in long-term memory to reconstruct degraded memory traces. Hence, the visuospatial domain may have some specificities such that visuospatial items can be refreshed only if a stable long-term memory representation exists for these items. The current study aimed at directly testing this proposal by manipulating the quality of long-term memory representations through the familiarity effect in visuospatial working memory.

4.1.4 The current study

The aim of this series of experiments was to investigate whether long-term memory factor influences attentional refreshing of information in visuospatial working memory. This would allow to examine the relationship between long-term memory and visuospatial working memory. Through four experiments, we adopted the approach of Camos and colleagues (2019) but in the visuospatial domain, which would allow us to investigate the domain-generality of attentional refreshing by comparing the results in the visuospatial domain to former experiments in the verbal domain. Hence, we needed an equivalent of lexical frequency or lexicality manipulation implemented by Camos et al. (2019), but in the visuospatial domain.

The lexical frequency effect relies on the fact that we do not encounter all words at the same frequency in day-to-day life. By definition, items are more familiar when encountered more frequently. It has been thus hypothesised that the strength of a trace in semantic long-term memory depends on how frequently the memoranda have been encountered before having to maintain the said trace. To emulate this in the visuospatial domain, we used the visuospatial familiarity effect by presenting two pools of images varying in their degree of familiarity. For the highly familiar images (here after the "real" images), we used black-and-white drawings of real-life objects taken from the Snodgrass and Vanderwart (1980) image set. From this set, we chose the most familiar images, selecting those for which names had a lexical frequency of 20 or more per million as indexed by the Lexique3 French database (New, Pallier, Brysbaert & Ferrand, 2004). The low familiar images (here after the "non-real" images) were taken from Soldan, Zarah, Hilton and Stern (2008). They consisted of images created by features from different images taken from the Snodgrass and Vanderwart image set smoothly put together (Figure 1). The resulting images were then approximately as complex as the original images but were not representing day-to-day objects. In our view, the difference between the real images and the non-real images is similar to the difference between words and pseudo-words. Pseudo-words are meaningless items that follow the phonotactical structure of their language of origin. In the same way, the "non-real" images were created using features coming from real objects that are usually not visually mixed together.

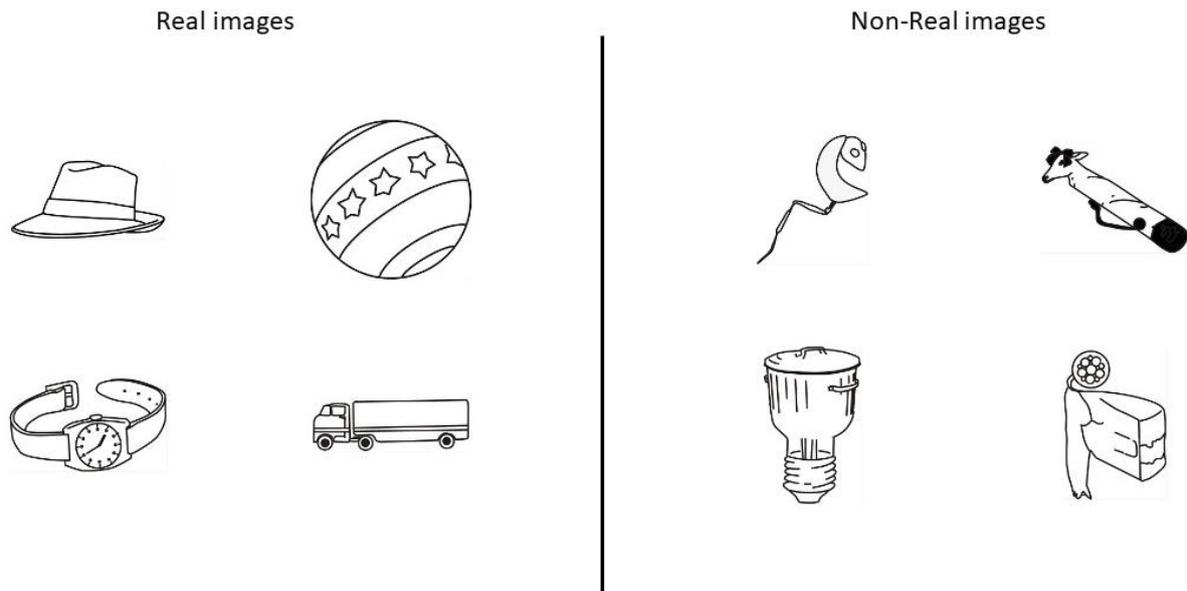


Figure 1: Examples of images used in the study.

We used the images described above in two series of two experiments. In Experiments 1A and 1B, we used a complex span task in which we manipulated orthogonally the familiarity of the images to maintain and the cognitive load of the concurrent task in a within-subject design. The aim was to investigate whether recall performance was better for familiar images compared to non-familiar ones and, most importantly, whether the cognitive load effect was lower (or even absent) for non-familiar images compared to familiar ones.

We expected that, if attentional refreshing is facilitated by a greater accessibility to semantic long-term memory representations, then image familiarity should interact with cognitive load effect, low-familiarity images yielding a smaller cognitive load effect than high-familiarity images. The cognitive load effect can be increased by reducing the time available to perform decision on distracting stimuli of the concurrent task or by increasing the number of stimuli to process in the processing phase of a complex span task (Barrouillet et al., 2007; Barrouillet & Camos, 2012). In Experiment 1A, the same number of digits was presented in the parity judgement phase of the complex span task, with either a long or short duration to judge each digit in the low and high cognitive load condition, respectively. In Experiment 1B, to create a greater contrast between the two cognitive load

conditions, we varied the number of digits presented during each processing episode, while keeping the interval duration constant. This allowed for a greater cognitive load difference between both conditions.

In Experiments 2A and 2B, we implemented a slightly different complex span task, designed to investigate how the response times to the concurrent task would be modulated by the image familiarity as well as the number of images to maintain. We expected that, if the real images are better refreshed than the non-real images, then the memory load effect on response times should be more pronounced for the non-real images compared to the real ones. This would result in a statistical interaction between memory load and image familiarity. Contrary to Camos et al. (2019), we choose to use the complex-span paradigm instead of the Brown-Peterson paradigm to evaluate response times to the concurrent task in regard to the memory load. Indeed, the complex span paradigm gave a higher number of data points per experimental cell during one session than the Brown-Peterson task, augmenting the reliability of our outcome variables. This is due to the fact that, in the Brown-Peterson paradigm, only one memory load condition is manipulated per trial. By contrast, one trial of the complex-span paradigm induces several memory loads along the trial, and response times to the concurrent task can be sampled after the presentation of each additional memoranda.

Since our theoretical question focused on the presence or the absence of statistical interactions, we used Bayesian statistics throughout our analyses. This statistical approach can give evidence for or against an effect. In the Bayesian approach, the resulting statistic is a number comprised between 0 and positive infinity. This number can be interpreted straightforwardly as how many times a model explains the data better (or worse) than another model. For example, when compared to the null model, a model with a Bayes factor of 10 implies that this model explains the data 10 times better than the null model. The Bayes factor can be calculated at the model level or at the variable level. When calculated at the variable level, it is

called the Bayes factor for the inclusion (or not) of a factor, which is the Bayes factor of the model with the variable of interest divided by the Bayes factor of the model without this variable of interest. The Bayes factor for (or against) the inclusion of a variable can also be interpreted straightforwardly as how much evidence there is in the data to include (or exclude) the variable. All analyses in this paper were performed with the BayesFactor package (Morey & Rouder, 2018) in R (R Core Team, 2020), with default settings.

4.1.5 Experiment 1A

In Experiment 1A, participants had to maintain the real and non-real images described earlier while performing a parity judgement task. We manipulated orthogonally the familiarity of images to maintain to the cognitive load of the concurrent task. We selected the parity judgement task as concurrent task because it requires participants to make a response selection and ensures that general attention is involved. This task is also known for impacting attentional refreshing (e.g., Barrouillet et al., 2007). To maximize the cognitive load difference between conditions, the same number of digits would appear slowly in the low cognitive load condition (one digit every 2000ms) and at twice this pace in the high cognitive load condition (one digit every 1000ms), which reduces the availability of attention for maintenance purpose. Participants were also under articulatory suppression during image presentation and the concurrent parity judgment task to minimize the possibility of recoding verbally and rehearsing the presented information. The method was particularly important as it can be assumed that the real images are easier to recode into verbal code than the non-real images. If participants were not under articulatory suppression, it would have induced a difference in coding and maintenance between the two conditions of image familiarity. We predicted that, if familiar images are refreshed more efficiently than less familiar images, then the cognitive load effect should be bigger with highly familiar items compared to less

familiar items, because the former would benefit more from the availability of attention during the concurrent task to be refreshed than the latter.

4.1.5.2 Method

Participants

Forty students from the University of Fribourg and Geneva (38 women, mean age: 20.8 ± 2.0 years) participated in this experiment. They were remunerated with partial course credits or cinema tickets. Every participant read and signed an informed consent form. Ethical approval was given by IRB of both the University of Fribourg and University of Geneva.

Material

The memory items for the familiar condition consisted of 84 images of real objects taken from the Snodgrass and Vanderwart (1980) database. We selected images related to concepts with high lexical frequency nouns (mean frequency = 107.9 ± 136.3 per million, range: 20 – 788, taken from the Lexique3 French words database; New, Pallier, Brysbaert & Ferrand, 2004) to ensure that the images represented well known objects. For example, images could represent an airplane, a finger, a pan (Figure 1). The stimuli for the unfamiliar condition comprised 84 non-real black and white images that were taken from Soldan et al. (2008). Each image was composed of features coming from different images from the Snodgrass and Vanderwart's image set fused together to create nonsensical objects. The complete set of images as well as the raw data from all the experiments presented in this study can be found on the following OSF page:

https://osf.io/xdawz/?view_only=9388c7ef0fa4430181876a955df2088c.

Procedure

This experiment used a complex span paradigm (Figure 2) in which stimuli to remember were the real and the non-real images described earlier. The concurrent task was a parity judgement task with digits ranging from 1 to 9, 5 being excluded for

having the same number of odd and even digits. The experiment was divided in two blocks, one for the real images and the other for the non-real images. The order of the blocks was counterbalanced across participants. Each block consisted of 42 trials. An increasing length procedure was used, from 1 to 7 memory items (length 1 to 7). Participants were informed of the change of lengths by a screen mentioning the new length. Each block started with six length-1 trials, three in each cognitive load condition, the order of which was randomized for each participant. We also implemented a stop-rule in each block. If the participant succeeded in at least one trial from a given length, the six trials from the next length were presented.

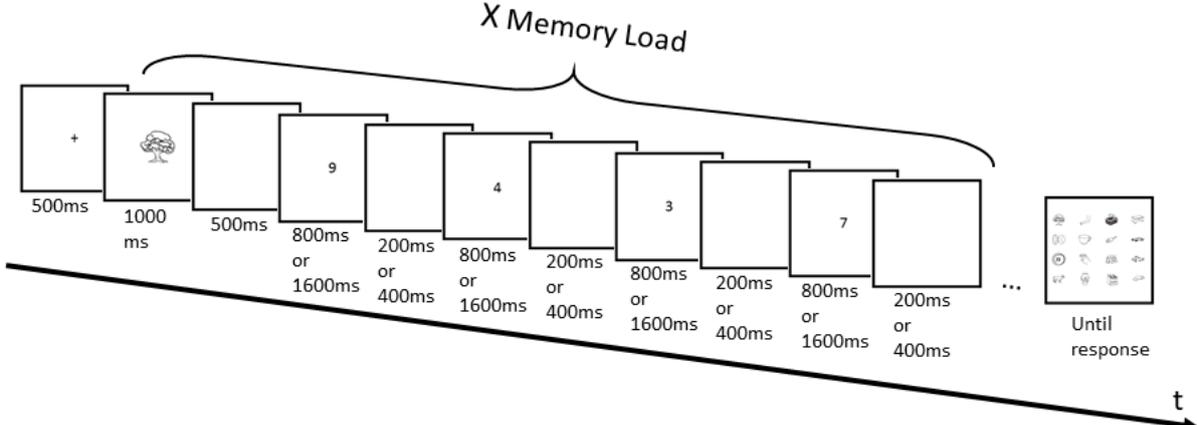


Figure 2: Illustration of the successive events in a complex-span task trial in Experiment 1A. The presentation of a memory image and the phase of parity judgments are repeated for each image presented in the trial.

Participants were seated at approximately 40 cm from the screen. At the start of every trial, the pace of the concurrent task was given: "rapide" or "lent" (fast and slow in French, respectively) and would remain on screen until the participant pressed the space bar to start the trial. Then, a fixation cross appeared in the center of the screen for 500ms, followed by the first image to memorize, which stayed on screen for 1s. The image was presented in the center of the screen and measured approximately 10cm by 6cm. After the image presentation and a 500ms blank screen, the parity judgment task started. The participant had to decide whether four sequentially presented digits were odd or even by pressing the left or right mouse button, respectively. In the high cognitive load condition, digits appeared for 800ms

with a 200ms blank screen after each digit. In the low cognitive load condition, digits stayed on screen for 1600ms with a 400ms blank screen. The parity judgment phase lasted then 8 seconds and 4 seconds in the low and high cognitive load conditions, respectively. The digits were randomly selected without replacement for each parity judgment task phase. At the end of a trial, participants were presented with an array of 16 images and had to click on the images that were presented in the order of presentation, using the left mouse button.

We chose an array of 16 images because it was a good balance between having a low chance level but still having the images being not too small on screen during the response phase. Participants had to click on the same number of images as the length of the trial and could not correct themselves after they had clicked on an image. They had no time limit to answer. The presented images were randomly selected from the pool of images used in the current block (real or non-real). During the response phase, the presented images were intermixed with non-target images coming from the same image pool. Since there were always 16 images presented simultaneously during the response phase, the number of non-target images was always equal to 16 minus the number of presented items in the current trial. The non-target images were not used in the trial preceding and following the current trial.

In the response phase, the images were 3cm by 2cm in size, and they were presented on a 4x4 matrix centered in the middle of the screen. Each image was separated by 2cm horizontally and vertically to any adjacent images. After the participant had given their answer, the screen provided information for the next trial. The choice of images in each trial was randomized for each participant, with the constraint that each image appeared a maximum of two times as memory item (once in each cognitive load condition) and six times as non-target during the response phase.

To minimize verbal recoding and the use of subvocal rehearsal, participants had to say "Ba-Bi-Boo" out loud from the start of each trial (i.e., when they pressed

space bar to launch the trial) until the response phase, where they could stop. They had to start again at the beginning of the next trial. “Ba-Bi-Boo” was written on the top of the screen from the beginning of the trial until the response phase to remind them to repeat it. Before the experimental trials, participants had 4 training trials, two length-2 and two length-3, with one training trial for each cognitive load in each length. Both image types were used for each list length during these training trials. The images used in the training trials were not used in any experimental trials. Before the training trials, participants underwent a training to the parity judgment task, where they had to sort 12 digits in each cognitive load condition, starting from the low cognitive load condition. The experimenter stayed with the participants during the training phase, but left them alone during the experimental trials while monitoring the compliance of the articulatory suppression across the door.

4.1.5.3 Results

First, we assessed participants’ performance to the parity judgment task to ensure that participants followed the instructions during the task. We computed the rate of correct parity judgements across the whole experiment. Participants with less than 70% of correct responses across all trials were discarded from analysis. This led to discarding data from three participants. We then used this ratio of correct parity judgements as dependent variable in a 2 (image familiarity: real vs non-real images) x 2 (cognitive load: high vs low) Bayesian repeated measure ANOVA, with the subject number as the repeated-measure aggregator and its prior set as “nuisance”. We used the default prior (medium) for the other factors. The best model included only evidence for a simple effect of the cognitive load ($BF_{inclusion} = 1.1 \times 10^{33}$). Evidence was against an effect of the image familiarity ($BF_{exclusion} = 4.0$) and against the interaction between image familiarity and cognitive load ($BF_{exclusion} = 3.2$). Although processing performance was better in the low cognitive load ($94\% \pm 3\%$) compared to the high cognitive load condition ($83\% \pm 6\%$), performance was high overall, ensuring a good reliance on the instructions.

We then calculated the span for each participant in each experimental cell. The span was calculated as follows: each trial in which all images were correctly recalled in the correct position yielded 1 point. Thus, as soon as one error was made, the trial yielded 0 point. Then, for each experimental cell, points were summed and divided by the number of trials per list-length (3). We used this mean span as our dependent variable in a 2 (image familiarity) x 2 (cognitive load) Bayesian repeated measure ANOVA. As in every analysis in this study, the priors for the effects of interest were set as "medium" (default) and as "nuisance" for the within-subject aggregator. The best model was the one with the two simple effects of image familiarity and cognitive load with a $BF_{10} = 6.8 \times 10^6$ (Figure 3). We found very strong evidence for the image familiarity effect, with a lower mean span for the non-real images (mean = 3.1 ± 1.4) than for the real images (mean = 4.0 ± 1.4 ; $BF_{inclusion} = 6.1 \times 10^6$). Although in the expected direction, the evidence for an effect of the cognitive load manipulation was more ambiguous, with a $BF_{inclusion}$ of 2.3 in favor of a difference between the low (mean = 3.7 ± 1.5) and the high (mean = 3.4 ± 1.4) cognitive load conditions. This probably reflected the fact that evidence for the cognitive load effect was clear in real images ($BF_{10} = 12.5$) while evidence supported its absence in non-real images ($BF_{01} = 3.1$). Nevertheless, and contrary to what these latter effects may suggest, there was ambiguous evidence against the interaction factor, $BF_{inclusion} = 0.6$.

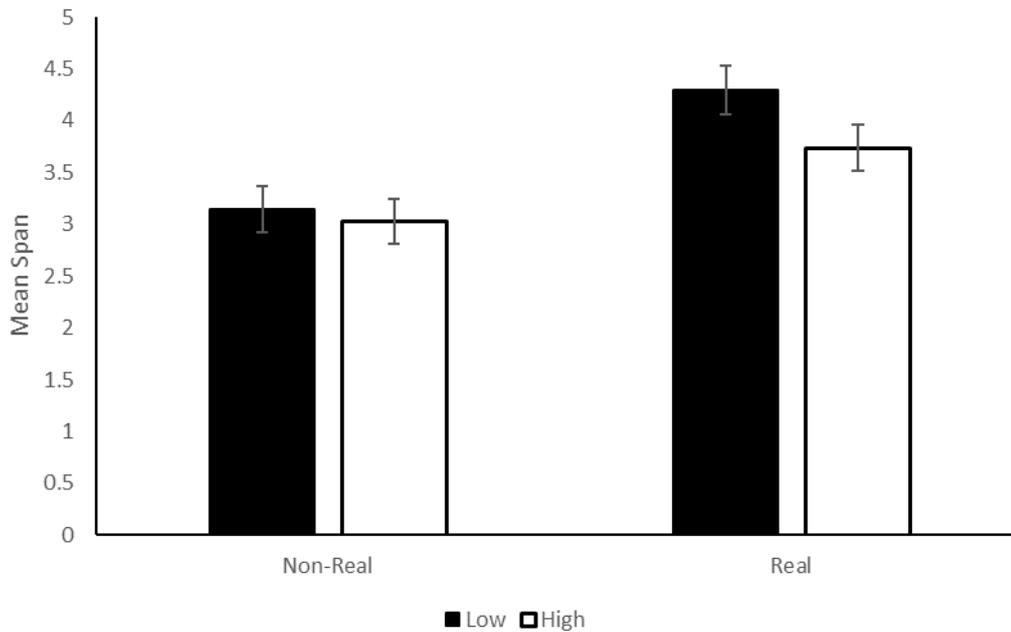


Figure 3: Mean span in Experiment 1A as a function of the type of memory images (Non-real or Real) and the cognitive load of the concurrent task (Low or High). Error bars correspond to standard error.

4.1.5.4 Discussion

In this first experiment, we observed a strong effect of image familiarity on memory performance in a working memory task, where real images yielded better memory performance than non-real images. We also found somewhat ambiguous evidence against the interaction between image familiarity and cognitive load effect. Even though the cognitive load effect was included in the best model of our memory data, the evidence for the inclusion of the cognitive load effect was rather weak. It is possible that the cognitive load manipulation we implemented was not effective enough to really impact recall performance. Hence, before drawing further conclusions on these findings, we implemented a stronger cognitive load manipulation in Experiment 1B.

4.1.6 Experiment 1B

Experiment 1B used the same paradigm as Experiment 1A with a few changes. First, the cognitive load manipulation was strengthened. Although we kept the same pace as in Experiment 1A for the low and high cognitive load, 8 digits were presented

in the high cognitive load condition and only 4 in the low cognitive load condition. This equalized the total duration (8s) of the parity judgment task across cognitive load conditions. We thus managed to manipulate the cognitive load of the secondary task without changing the total duration of each secondary task phase. We also divided each pool of image (Real and Non-Real) into two sublists. Each participant was attributed one sublist to the high cognitive load condition and the other to the low cognitive load condition, the attribution of which was counterbalanced across participants. This allowed us to use each image only once as a memorandum throughout the whole experiment, minimizing a possible training effect on specific images, and also to evaluate if one half of the images yielded better performance compared to the other half. An absence of such an effect of list would show that our results are not dependent on the actual images use in the experiment, and are thus more generalizable.

4.1.6.1 Method

Participants

Forty-one students from the University of Geneva (35 women, mean age: 22.2 ± 6.5 years) participated in this experiment. They were compensated with partial course credits. Every participant read and signed a form of consent. Ethical approval was given by IRB of the University of Geneva. None of them participated in Experiment 1A. One participant was excluded from the analysis because he did not follow the instructions.

Material

The same 84 images per image type as in Experiment 1A were used and divided into two sublists for each type of images. For each participant, one sublist was presented in the low cognitive load condition and the other in the high cognitive load condition. The association between sublist and cognitive load condition was

counterbalanced across participants. This allowed us to use every image only once as a memorandum and approximately 3 times as non-target in the response phase.

Procedure

This experiment followed the Experiment 1A procedure with a few differences. Each block consisted of 20 trials, presented in an increasing length structure. The list length varied from 2 to 6, with 4 trials per each list length and image type, two in each cognitive load condition. The pace of digits presentation in the parity judgment task was the same as in Experiment 1A, but 8 digits were presented in the high cognitive load, and 4 in the low cognitive load condition to increase the difference of cognitive load between the two conditions.

4.1.6.2 Results

As in Experiment 1A, we first assessed performance to the parity judgment task. We did not exclude any participants as they all exhibited more than 70% of correct response (mean = 89% \pm 6%). We then analyzed the ratio of correct parity judgement in a 2 (image familiarity) \times 2 (cognitive load) repeated measure Bayesian ANOVA. The best model to come out of this analysis only had a main effect of cognitive load manipulation ($BF_{10} = 4.1 \times 10^{27}$). We found evidence for an effect of the cognitive load manipulation ($BF_{inclusion} = 5.4 \times 10^{27}$) with better performance in the low cognitive load condition (95% \pm 5%) compared to the high cognitive load condition (86% \pm 7%), evidence against an effect of image familiarity ($BF_{exclusion} = 4.3$), and evidence against the interaction ($BF_{exclusion} = 5.2$).

Spans were computed for each participant and each experimental cell in the same way as in Experiment 1A. We first checked the absence of the sublists manipulation on the span score by performing an independent sample Bayesian t-test on the span score of participants that had one sublist compared to the other. We found evidence against an effect of which sublist was used ($BF_{exclusion} = 3.1$). We then turned to the analysis of the span score as a function of image familiarity and

cognitive load manipulation. A 2 (image familiarity) x 2 (cognitive load) repeated measure Bayesian ANOVA was performed with the calculated span as dependent variable (Figure 4). The same priors as in Experiment 1A were used for the analysis. The best model included the simple effects of image familiarity and cognitive load, $BF_{10} = 1.9 \times 10^{10}$, with better performance for the real images (mean = 4.1 ± 1.3) than for non-real images (3.2 ± 1.2 , $BF_{inclusion} = 3.2 \times 10^9$), and better performance in the low (3.8 ± 1.3) compared to the high cognitive load condition (3.3 ± 1.3 , $BF_{inclusion} = 19.6$). Finally, despite the fact that the strengthening of cognitive load manipulation was successful on impacting recall performance, there was evidence against the predicted interaction between cognitive load and image familiarity ($BF_{exclusion} = 4.2$).

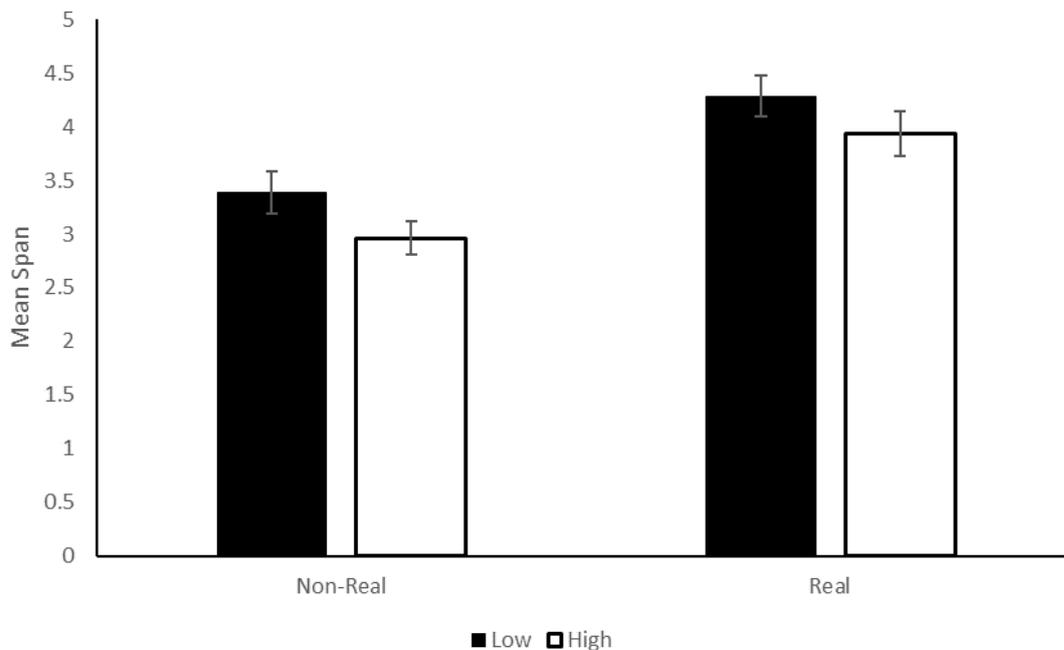


Figure 4: Mean span in Experiment 1B as a function of the type of images to maintain (Non-real or Real) and the cognitive load of the concurrent task (Low or High). Error bars correspond to standard error.

4.1.6.3 Discussion

In this second experiment, we replicated the findings of Experiment 1A but with a clearer outcome. We found more pronounced evidence for the effect of cognitive load as well as for the effect of image familiarity, but also against an interaction between both factors. Overall, Experiment 1B confirmed the absence of

the interaction of interest. This contradicts the hypothesis that attentional refreshing is supported by semantic long-term memory representations.

However, to ensure that our conclusion on the absence of an effect of semantic long-term memory on attentional refreshing was not limited to a single index of attentional refreshing, we extended our investigation of the image familiarity effect to another index of attentional refreshing, namely the memory load effect on response times in the concurrent task. As explained earlier, several studies showed that an increase in memory load induced a postponement of the response times to the concurrent task, which is interpreted as evidence for attentional refreshing. The use of response times allowed us to examine the effect of image familiarity on attentional refreshing in a more fine-grained measure than span. If the more familiar images are refreshed more efficiently than less familiar ones, we should expect a difference in the memory load effect on response times as a function of image familiarity. The more familiar images should elicit less postponement of the concurrent task than less familiar images (i.e., smaller memory load effect for the less familiar images compared to the more familiar ones). To test this hypothesis, we adapted the complex-span paradigm used in Experiments 1A and 1B to directly measure the response times to the parity judgement task in two experiments.

4.1.7 Experiments 2A and 2B

The aim of these two experiments was to investigate the effect of image familiarity and memory load on the response times to the parity judgment task in a complex-span paradigm. The hypothesis was that if the functioning of attentional refreshing is influenced by image familiarity, then we should detect an interaction between the memory load manipulation and the image familiarity manipulation on response times in parity judgements. Experiments 2A and 2B followed the same structure as Experiments 1A and 1B, with the same images to manipulate image familiarity. The only difference between Experiments 2A and 2B concerned the implementation of the concurrent task.

It is important to understand that to examine the memory load effect on response times, participants need to correctly recall the memory items to assure they were maintained in working memory, but participants also have to make correct judgments in the concurrent task. Hence, the way that the concurrent task is implemented can have an impact on participants' performance. We chose then to administer two variations of the same parity judgement task to check that our findings were not dependent on methodological choices. In Experiment 2A, participants had to judge the parity of a fixed number of digits (4 digits after each memory item), with no time limit on each parity judgement, while in Experiment 2B, the concurrent task between the presentation of memory items lasted for a fixed duration (5 seconds after each memory item), and participants had to judge as many digits as possible during this time. If the memory load effect on response times is larger in the less familiar image condition compared to the more familiar image condition, this would mean that the former is refreshed less quickly and less efficiently than the latter.

4.1.7.1 Method

Participants

In total, 57 participants were recruited and randomly assigned to one of the two experiments, 29 to Experiment 2A (26 women, mean age = 20.8 ± 2.1 years) and 28 to Experiment 2B (26 women, mean age = 21.6 ± 2.4 years). All were students from University of Fribourg and remunerated with partial course credits. Every participant read and signed a form of consent. Ethical approval was given by IRB of the University of Fribourg. None had participated in Experiments 1A or 1B. One participant from Experiment 2B was excluded from the analysis for having stopped the articulatory suppression.

Material and Procedure

The same general procedure as in Experiments 1A and 1B was applied, with changes in the implementation of the concurrent task, on the total number of trials and the memory load manipulation.

First, there were 2 versions (named A and B) of the parity judgment task, one per experiment. In Experiment 2A, the number of digits was fixed, with 4 digits presented sequentially. In Experiment 2B, the duration of the parity judgment phase between the presentation of the memory images was fixed, lasting 5s, during which participants had to judge as many digits as possible. Participants were instructed to perform the parity judgement task in such a way that, though aiming at responding as fast and as accurately as possible, they remembered all the memoranda in their order of presentation. In both experiments, the next digit appeared 50ms after the last one was judged, and the delay between image presentation and the beginning of the concurrent task were shortened to 300ms. Second, the list length varied from 2 to 4 images with 8 trials in each length for each type of images (i.e., 48 experimental trials in total). Hence, seventy-six images were selected from the Snodgrass and Vanderwart (1980) for the familiar images and from Soldan and colleagues (2008) for the less familiar images. Based on Experiments 1A and 1B, we expected that most of participants would succeed recall in at least two length-4 trials. All trials were presented in one single block, with the trial order randomized for each participant. Before each trial, the number of memory images was indicated. Items for each trial were selected randomly with the limitation that items were not used twice as memoranda for the same participant.

Before the experimental trials, participants did 4 training trials to accustom them to the task, two trials of length-2 and two trial of length-3. Both kinds of images were used during training, one for each list length. The version of the parity judgement task (A or B) was the same in the training and experimental trials. None of the images used in the training trials were used in the experimental trials.

Three different outcome variables were analyzed: recall performance, the first response time of every parity judgement phase (here-after: Initial-RT) and the mean of response times to all of the subsequent parity judgements of each parity judgment phase (here after: Subsequent-RT). To examine the effect of the memory load on the response times in the parity judgements, only trials with correct recall were included in the analyses (see Camos et al., 2019; Vergauwe et al., 2014). We split response times in Initial-RT and Subsequent-RT, because response times in the first parity judgement is contaminated with switching processes and consolidation process linked to image presentation, and thus are significantly longer than any subsequent response times (Camos et al., 2019; Jarrold et al., 2011; Vergauwe et al., 2014). Since we sampled the Initial-RT and Subsequent-RT as a function of memory load, multiple Initial-RT and Subsequent-RT values could be measured during a single trial. For example, in a length-4 trials, we had four measures for both response times: The Initial-RT following each image presentation (i.e., RT to the first digit that followed memory item presentation) as well as the mean of the Subsequent-RTs following each image presentation.

4.1.7.4 Results

As in Experiments 1A and 1B, we used the ratio of correct parity judgements as exclusion criterion. Since our analysis was based on correct parity judgments, we chose an exclusion threshold higher than in Experiments 1A and 1B. Participants should have more than 80% of correct parity judgements to keep their data. No participants were excluded based on this criterion in Experiment 2A and one in Experiment 2B. While participants judged the parity of the 4 digits presented after each memory image in Experiment 2A, they judged on average 6.2 ± 1.2 digits (range: 3-10) in Experiment 2B, with no difference in the number of processed digits between conditions of image familiarity (Bayesian t-test on the number of processed digit between high and low familiarity conditions, $BF_{\text{exclusion}} = 5.5$). In both experiments, participants did their best to follow instructions and their rate of correct

parity judgements was high (97%, 96%, and 96% in Exp. 2A, and 95%, 95%, and 94% in Exp. 2B, for lengths 2 to 4, respectively).

Before analyzing response times to the parity judgement task, we first examined recall performance. Because no stop rule was implemented, recall performance was scored via the partial credit unit (PCU). This index is the mean proportion of images correctly recalled in the correct position in each trial. The resulting score has been shown to be a better index of recall performance than span in working memory tasks (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005). We then did a two-sided paired Bayesian t-test on PCU score between the low and high visuospatial familiarity condition. In both experiments, this analysis showed a difference between high and low familiarity images. Trials with non-real images (PCU= 0.61 ± 0.24 and 0.63 ± 0.20 in Experiments 2A and 2B, respectively) yielded worse recall performance than trials with real images (0.74 ± 0.19 and 0.78 ± 0.16), $BF_{10} = 333$ and 1048 , for Experiments 2A and 2B, respectively, replicating the visuospatial familiarity effect observed in Experiments 1A and 1B.

Next, we analyzed the response times in the concurrent parity judgment task. To investigate the effect of memory load as well as of visuospatial familiarity on response times, only trials with perfect recall were kept, leading to 49% and 47% of the trials being discarded in Experiments 2A and 2B, respectively. The mean Initial-RT and mean Subsequent-RT were then computed for each participant in each of the 8 experimental cells (2 image familiarity: Real or Non-Real x 4 memory load: 1 to 4 images). To ensure reliability of our response time measures, experimental cells in which less than two correct trials were available for computing the mean were discarded. This led to discarding 42 experimental cells out of the 232 (8 x 29 participants) in Experiment 2A (3 from the 1-image and the 2-image conditions, 12 from the 3-image and 33 from the 4-image conditions) and 29 out of the 208 cells (8 x 26 participants) in Experiment 2B (6 from the 3-image and 23 from the 4-image

conditions). Initial-RTs and Subsequent-RTs were then analyzed in two separate repeated measure Bayesian ANOVAs.

Before averaging RTs across the subsequent positions (i.e., across all digits presented within a processing episode, except the first digit), we analyzed the effect of position of digits within a processing episode (see supplementary material). Results of this analysis showed that, in both experiments, there was no effect of digit position (from position 2 onward) on response time to the concurrent task. Subsequent-RT from the same parity phase and from the same participant could thus be pooled together.

We thus averaged RTs across subsequent positions (i.e., all except the first position) and analyzed them separately for Experiments 2A and 2B in two 2 (Image familiarity: Real or Non-Real) x 4 (Memory load: 1 to 4) repeated measure Bayesian ANOVAs, one per experiment. Both analyses yielded the same pattern of results (Figure 5). The best models included only a main effect of the memory load ($BF_{10} = 4.8 \times 10^4$ in Experiment 2A and $BF_{10} = 3.9 \times 10^8$ in Experiment 2B). There was overwhelming evidence for an increase in response times as a function of the memory load (609 ms \pm 101ms, 659ms \pm 131ms, 681 \pm 122, and 704ms \pm 190m for 1 to 4 images in Experiment 2A; 617 ms \pm 122ms, 658ms \pm 151ms, 680 \pm 163, and 741ms \pm 175ms in Experiment 2B). We also found evidence against an effect of image familiarity (652ms \pm 118ms and 663ms \pm 152ms for non-real and real images in Experiment 2A; 655ms \pm 154ms, and 674ms \pm 156ms in Experiment 2B),

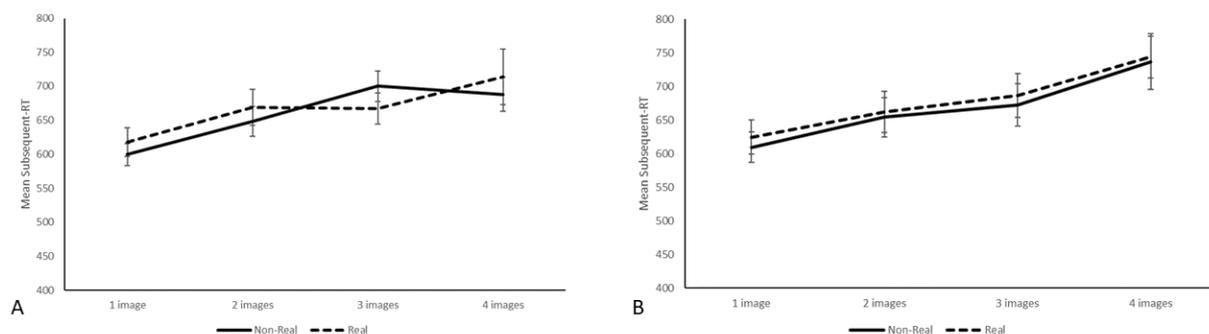


Figure 5: Mean Subsequent-RT (in ms) in Experiments 2A (left) and 2B (right) as a function of the memory load (1 to 4) and the image familiarity (non-real or real). Error bars correspond to standard error.

$BF_{\text{exclusion}} = 5.5$ and 5.2 in Experiments 2A and 2B, respectively, and against the interaction between image familiarity and memory load, $BF_{\text{exclusion}} = 7.2$ and 9.7 , respectively.

To assess the speed of refreshing, we computed the slope of a linear regression on the mean Subsequent-RTs as a function of the memory load, separately for the Real and the Non-Real images in Experiments 2A and 2B. In Experiment 2A, the linear regression showed an increase of 29ms per new image held in working memory for the Real images ($R^2 = 0.89$) and 32ms per new image for the Non-real images ($R^2 = 0.81$). In Experiment 2B, the linear regression showed an increase of 38ms per new image for the Real images ($R^2 = 0.97$) and an increase of 40ms per new image for the Non-real images ($R^2 = 0.96$).

The analysis of the Initial-RTs also yielded a similar pattern in both experiments (Figure 6). The best models in the two experiments were the full model with both the main effect of memory load and image familiarity in addition to the interaction between these two factors ($BF_{10} = 3.8$ and $BF_{10} = 1.1 \times 10^{10}$ in Experiments 2A and 2B, respectively). However, in both experiments, the full model was only ambiguously better than the model that included the effect of the memory load only, with Bayes factors for the memory only model that are similar to those observed for the full model ($BF_{10} = 3.7$ for the memory only model against the null model, and $BF_{10} = 5.2 \times 10^9$ for the memory only model against the null model in Experiments 2A and 2B, respectively).

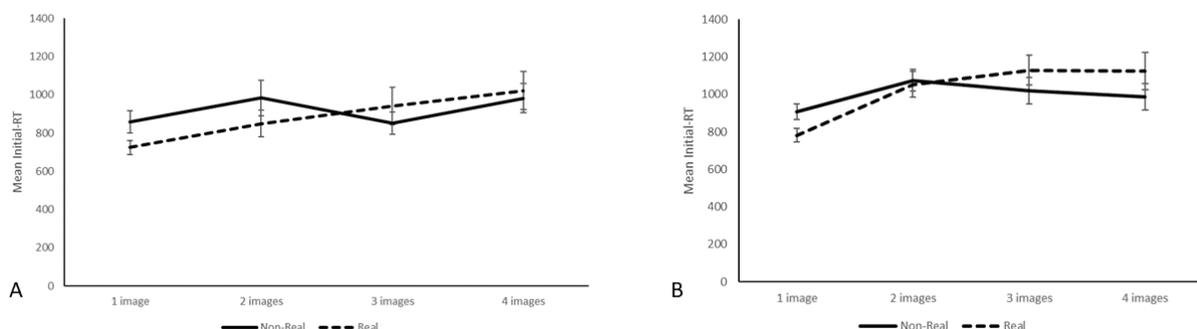


Figure 6: Mean Initial-RT (in ms) in Experiments 2A (left) and 2B (right) as a function of the memory load (1 to 4) and the image familiarity (non-real or real). Error bars correspond to standard error.

We compared the full model and the memory load only model separately for each experiment. The full model explained our data only 1.1 times better in Experiment 2A and only 2.3 times better in Experiment 2B, compared to the memory load only model.

4.1.7.5 Discussion

Despite differences in the implementation of the concurrent task, both experiments yielded the same pattern of results. First, we replicated the visuospatial familiarity effect in memory performance. Non-real images yielded worse recall performance than real images. Furthermore, as expected, the analysis on the Subsequent-RTs showed that the response times to the concurrent task was influenced by the memory load. Indeed, response times were postponed by approximately 30 - 40ms for each additional image to maintain. This is in line with previous findings estimating the speed of refreshing around 40ms (Camos et al., 2019; Jarrold et al., 2011; Vergauwe et al., 2014). Finally, we also found evidence against the interaction between image familiarity and memory load on the Subsequent-RTs, which contradicts our hypothesis of an involvement of semantic long-term memory in attentional refreshing.

The analysis on the Initial-RTs showed an effect of the memory load on this outcome variable. Initial-RT were slower the more images were maintained in working memory. This could be evidence for a consolidation process happening in between item encoding and maintenance. However, if Initial-RTs was only due to the consolidation of a newly presented image, it is striking that these Initial-RTs increased across serial positions, with longer Initial-RTs after the third image than after the first or second image, for example. In the General Discussion, we will elaborate further on this issue.

4.1.8 General Discussion

In a series of four experiments, we investigated the hypothesis that the functioning of attentional refreshing relies on semantic long-term memory representations in the visuospatial domain of working memory. More specifically, we hypothesized that visuospatial items that are more familiar (and thus better represented in semantic long-term memory) would be refreshed more efficiently and more quickly than less familiar items. In Experiments 1A and 1B, we tested this hypothesis by manipulating orthogonally item familiarity and refreshing opportunity in a complex-span task and evaluated their impact on recall performance. The prediction was that if highly familiar items are refreshed more efficiently than less familiar items, then the cognitive load effect should be more pronounced for highly familiar items compared to less familiar items, and this would result in the presence of a statistical interaction between the cognitive load and the familiarity manipulation on recall performance. In Experiment 1A, we found evidence for an effect of the familiarity manipulation and ambiguous evidence for the cognitive load effect, as well as ambiguous evidence against the interaction. In Experiment 1B, we strengthened the cognitive load manipulation and, this time, found persuasive evidence for an effect of the cognitive load manipulation as well as for the presence of an effect of the familiarity manipulation. However, convincing evidence was gathered against the interaction. Due to the evidence against the interaction between item familiarity and cognitive load in both experiments, our results are inconsistent with the hypothesis that attentional refreshing functioning is influenced by semantic long-term memory factors.

To examine this hypothesis with a more fine-grained measure, Experiments 2A and 2B investigated the effect of familiarity manipulation on another index of attentional refreshing: the memory load effect on concurrent task response times. If more familiar items are refreshed more efficiently than less familiar items, then the memory load effect on the response times to the concurrent task should be larger for

less familiar items compared to highly familiar items. This would result in the presence of an interaction between item familiarity and memory load manipulation. Moreover, the slope-related response times to memory load and indexing refreshing speed should be larger in less familiar items. We thus manipulated orthogonally the memory load and the familiarity of visuospatial items in two complex-span tasks, which differed slightly in their operationalization of the concurrent task. Experiments 2A and 2B yielded the same pattern of results, with evidence for the presence of a memory load effect on Subsequent-RTs to the concurrent task, but evidence against an effect of image familiarity and against the predicted interaction, the speed of refreshing being similar in familiar and less familiar images. Thus, thru four experiments, we gathered a very consistent pattern of findings, with convincing evidence for an effect of image familiarity and cognitive load manipulation on recall performance, as well as evidence for an effect of memory load manipulation on response times to the concurrent task. However, evidence points systematically against the hypothesized interactions. In the following, we will discuss the functioning of attentional refreshing in light of these findings and possible limitations to the functioning of attentional refreshing. We will also discuss the relationships between semantic long-term memory and working memory.

4.1.8.1 On the functioning of attentional refreshing

Several models of attentional refreshing have been put forward (see Camos et al., 2018, for a review). Although the models differ on several key aspects of attentional refreshing functioning, they tend to agree that attentional refreshing is domain-general. Comparing our results to Camos et al. (2019) gives support to this assumption. Our patterns of results in the visuospatial domain are very similar to the patterns reported by Camos et al. (2019) in the verbal domain.

Firstly, we replicated the cognitive load effect with visuospatial material and a recognition paradigm. Recognition performance was worse when participants had more digits to judge or less time to do so during the parity phase, compared to when

they had fewer digits/more time to judge. The effect was clearly present in Experiment 1B, but more ambiguous in Experiment 1A. This discrepancy was linked to the fact that the cognitive load manipulation was probably not strong enough in Experiment 1A to elicit a pervasive cognitive load effect, but it was clearly the case in Experiment 1B, in which the manipulation of cognitive load was strengthened. Our results also show that the findings from Camos et al. (2019) are not limited to one testing procedure. Compared to Camos et al. (2019) study, which used a written recall procedure, our experiment used a serial recognition paradigm. Although the pattern of results was similar in both studies, the cognitive load effect was less evident in our experiments compared to the same manipulation in the verbal domain, as indexed by the size of the Bayes factors. Since recognition does not demand a complete reconstruction of the mnemonics traces, less stable working memory representation can be recognized more easily than directly recalled. Participants could then have not actively maintained visuospatial items. This might be true for Experiment 1A, as the evidence for the effect of the cognitive load manipulation was ambiguous. However, when cognitive load was manipulated more strongly, the cognitive load effect was convincingly present in our data. This shows that participants actively maintained the images, at least in Experiment 1B (compared to Experiment 1A). We otherwise would not have detected a cognitive load effect on the recall performance. All in all, the results from Experiments 1A and 1B were in line with previous studies that found a cognitive load effect on recall performance with verbal as well as visuospatial material (Barrouillet et al., 2007, Barrouillet & Camos, 2012; Vergauwe et al., 2009, 2012, 2014).

Secondly, we also detected the memory load effect on response time to the concurrent task described in the literature with visuospatial material (Camos et al., 2019; Jarrold et al., 2011; Vergauwe et al., 2014). Interestingly, although our implementation of response time measures to the concurrent task was done in a complex-span paradigm instead of the Brown-Peterson task used by Vergauwe et al. (2014) and Camos et al. (2019), we found very similar pattern of results with a time postponement dependent on the memory load. Moreover, our estimate of the

refreshing rate at around 30-50ms per new item held in working memory was akin to previous estimates (35-45ms in Camos et al., 2019; 41ms in Jarrold et al., 2011; 28-40ms in Vergauwe et al., 2014). Such a memory load effect is congruent with the idea that attentional refreshing functions via a sequential scanning of all items held in working memory (Vergauwe & Cowan, 2015). Each item held in working memory before a concurrent task starts is refreshed one after the other in a sequential manner, before participants process the presented distractors on which they perform the concurrent task. Besides strengthening this account of the functioning of attentional refreshing, our results extend the domain-generalty of the memory load effect on the response times to a concurrent task. Indeed, the present study observed this effect in a paradigm in which it was never tested before (i.e., complex span task with serial recognition) and for visuospatial material that has never been introduced as memory items in a working memory task before. Because this postponement is considered as an index of attentional refreshing speed, the similarity of slope estimates indicates that visuospatial and verbal items are refreshed at a similar speed. This would be indicative that the same mechanism, probably refreshing, is at play in both domains. Thus, taken together with the results from Camos et al. (2019), our results reinforce the idea that attentional refreshing is a domain-general mechanism that works in the same way for verbal and visuospatial material.

In addition to the cognitive load and the memory load effect, our results also show that image familiarity has an impact on recall performance. In all our four experiments, we found conclusive evidence that non-familiar images yield worse recall performance than familiar images. This difference in performance shows that our manipulation of semantic long-term memory in the visuospatial domain was effective. Furthermore, the presence of this familiarity effect was also essential to test our hypothesis. Without such an effect on recall performance, no conclusion could be made on the presence or absence of the interaction between item familiarity and attentional refreshing manipulation (the cognitive load and the memory load manipulation). Despite the presence of the familiarity effect, we found evidence

against the interactions of interest in all four experiments. In Experiments 1A and 1B, although recall performance was impacted by the cognitive load manipulation and the familiarity manipulation, the cognitive load effect was similar in both type of images. This indicates that non-familiar images are refreshed as efficiently as familiar ones.

This conclusion is supported by the results of Experiments 2A and 2B, in which we found evidence against an effect of image familiarity and its interaction with the memory load effect on response times to the concurrent task. In these two experiments, there was only evidence for an effect of the memory load manipulation. The absence of interactions in Experiments 2A and 2B shows that familiar and non-familiar images were refreshed at the same speed. Taken together, results from our experiments show that image familiarity does not influence the functioning of attentional refreshing, even though it impacts recall performance. This contradicts the hypothesis that attentional refreshing functions via the involvement of semantic long-term memory.

4.1.8.2 Refreshable and non-refreshable material

To summarize, our results are congruent with the domain-general assumption of attentional refreshing and with previous studies that showed an absence of semantic long-term memory effect on attentional refreshing. However, the question remains as to why some specific visuospatial items are not influenced by attentional refreshing manipulation (Ricker & Vergauwe, 2020; Vergauwe et al., 2014). In their study, Ricker and Vergauwe (2020) failed to find cognitive load effect with some specific visuospatial material (i.e., position around a circle in an angle reproduction task) and put forward several explanations for the absence of this effect. They hypothesized that it could be due to the memory material itself, as an angle reproduction task had never been used before in any study with a cognitive load manipulation. It would suppose that specifics about position around a circle would prevent it to be refreshed. However, this does not seem to hold, as other types of

visuospatial memoranda have yielded an absence of memory load effect (Vergauwe et al., 2014) or seem to not be actively refreshable (Ricker & Cowan, 2010). In the current literature, other memoranda for which an effect of attentional refreshing manipulation was not observed are unconventional characters (Ricker & Cowan, 2010), and different fonts of the same letter (Vergauwe, 2014). Although the three kinds of memoranda pertain to the same domain, it is not clear how they relate to each other or what are their common features that drive the similar absence of effect.

As suggested by the authors finding an absence of cognitive load effect (Ricker & Vergauwe, 2020; Vergauwe et al. 2014), we entertained in the present study the idea that differences in semantic long-term memory mnemonic traces could be the basis of this discrepancy, but we failed to find support for this hypothesis. Another possible explanation relies on the possibility to verbalize the memoranda. One could argue that fonts of the same letter and unconventional characters are difficult to verbalize without a proper training, and that attentional refreshing relies on some sort of verbal recoding. However, this is not congruent with the results from Ricker and Vergauwe (2020). In Experiment 2, they used “canonical” positions around the circle instead of random continuous position, and still found no effect of the cognitive load manipulation, even though the actual position could easily be verbally recoded as “up”, “down”, “left” or “right”.

One last possible explanation relies on the relationship between consolidation and refreshing in working memory. Consolidation in working memory is a process that transforms fleeting sensory traces into more stable representations in working memory (Bayliss, Bogdanovs & Jarrold, 2015; Ricker & Cowan, 2014; Schrijver & Barrouillet, 2017; See Ricker et al., 2018, for a review). Ricker and Vergauwe (2020) argued that it could be possible that only well-enough consolidated traces can be refreshed, and that low-level visuospatial features are not easily consolidated. As mentioned in discussion of Experiments 2A and 2B, it is rather striking that the mere

consolidation of a newly presented image leads to a memory load effect on Initial-RTs.

Alternatively, it can be envisioned that the addition of a new image requires a reconfiguration of the output program. This could account for an increased Initial-RT throughout serial position because, due to the output program becoming more and more complex, it needs more time to be reconfigured (see Jones & Macken, 2018; Joseph & Morey, in press; Myers, Stokes & Nobre, 2017; Stokes, 2015). Overall, future studies should aim at finding which kind of material can be influenced or not by attentional refreshing manipulation, as it could give important limitations to the functioning of attentional refreshing and, thus, on active maintenance in working memory as a whole.

4.1.8.3 Relationships between visuospatial working memory and semantic long-term memory

Our results go against the idea that semantic long-term memory is directly involved in the functioning of attentional refreshing. However, since we did find a familiarity effect on recall performance, our results are congruent with the idea that semantic long-term memory has an impact on working memory functioning as a whole. The familiarity effect could originate from several other processes in working memory than maintenance period, namely during the encoding and/or the recall phase (see Thorn, Frankish, & Gathercole, 2009, for a review).

Regarding encoding processes, it could be that more familiar items are encoded more strongly than less familiar ones, resulting in better recall performance for more familiar items. However, our results in Experiment 2A and 2B on Initial-RTs showed that they are only influenced by the memory load and not by image familiarity. This is congruent with the idea that encoding processes happen at the same speed for familiar and unfamiliar items, but does not give information about the strength of this encoding. Nonetheless, future studies should aim at directly

investigating this hypothesis, as our experiments were not designed to directly investigate it.

Interestingly, the Initial-RTs pattern of results once again replicates the results from Camos et al. (2019) in the verbal domain, which is also consistent with the supposed domain-generalty of attentional refreshing. However, we do not consider this as definitive evidence. Initial-RTs were not our primary variable of interest, and the experiments were not designed to directly test the hypothesis that more familiar items are better encoded than less familiar items. Future experiments should aim at investigating this possibility.

Alternatively (or in addition to the possible effect at encoding described above), item familiarity could have an impact at the recall stage through the redintegration process (Hulme et al., 1997; Schweickert, 1993). This hypothesis supposes that long-term memory is used to reconstruct mnemonic traces in working memory at recall. Thus, items that are easier to retrieve from semantic long-term memory (i.e., high familiarity items) would have a greater probability to be redintegrated than harder-to-retrieve items (i.e., low-familiarity items). To conclude, our study goes against the idea that attentional refreshing, a major working memory maintenance mechanism relies on long-term memory, which tends to support the separation between working memory and long-term memory. Future studies should aim at disentangling the different possible loci for the familiarity effect, and more generally for the long-term memory effects that impact recall at short term.

4.1.9 Conclusion

Although we repeatedly found a familiarity effect on recall performance in our study, our data are not congruent with the hypothesis that more familiar visuospatial items are refreshed more efficiently than less familiar items. This shows that familiarity effect in working memory does not emerge from the maintenance phase. In four experiments, we found evidence against the hypothesis that attentional refreshing uses semantic long-term memory to reconstruct traces maintained in

working memory during the retention interval. The pattern of results corroborates what was already found in the verbal domain. Moreover, our results were also not congruent with the idea that, in the visuospatial domain, information can only be refreshed if a stable semantic long-term memory traces exist. Future studies should aim at evaluating possible boundary conditions to the functioning of attentional refreshing, as attentional maintenance process is a central part of working memory.

4.1.10 Supplementary Material

We analysed the response times (RTs) to the concurrent task as a function of digit position within a parity phase. This included the second, third and fourth digits in Experiment 2A and the second to eighth digits in Experiment 2B³. A Bayesian ANOVA was performed on RTs to the concurrent task as a function of digit position within a parity phase (2 to 4 in Exp. 2A, and 2 to 8 in Exp. 2B), image type (Real or Non-Real) and list length (2 to 4 memory items). In Experiment 2A, the best model was the null model, with evidence against an effect of digit position ($BF_{\text{exclusion}} = 4.1$), against the interaction between digit position and image type ($BF_{\text{exclusion}} = 19.3$), and against the interaction between digit position and list length ($BF_{\text{exclusion}} = 1.9$). We concluded that digits in position 2 to 4 could be pooled together in Experiment 2A for further analyses.

Experiment 2B showed a different pattern, with the best model including an effect of list length and an effect of digit position ($BF_{10} = 3.0 \times 10^{19}$). RTs to the concurrent task decreased with each new digit to judge within a parity phase (651ms \pm 158ms, 649ms \pm 131ms, 643ms \pm 116ms, 620ms \pm 101ms, 590ms \pm 63ms, 552ms \pm 51ms, 535ms \pm 69ms, for digit positions 2 to 8 respectively, $BF_{\text{inclusion}} = 3.5 \times 10^{17}$). However, this effect is likely due to the way we pooled the RTs: participants who judged eight digits during a parity phase (limited to 5 s) necessarily did it faster than participants who only sorted 4 or 5 digits. When averaged together, only the response times from “faster” participants (participants that judged more digits) are taken when averaging on the later digit positions. This is evidenced by two analyses: first, when only the digits in positions 2 to 4 were taken into account (maximum number of digits judged by all participants), the repeated Bayesian ANOVA on response time to the concurrent task as a function of digit position within a parity phase (2 to 4 in Exp. 2A and Exp. 2B), image type (Real or Non-Real) and list length (2 to 4 memory items)

³ In Experiment 2B, we included digits up to the eighth position, because 73% of our participants performed correctly at least one trial with a minimum of 8 digits sorted during a single parity phase.

showed only an effect of list length ($BF_{10} = 1.7 \times 10^3$), with evidence against an effect of digit position ($BF_{\text{exclusion}} = 32$), and evidence against an effect of image type ($BF_{\text{exclusion}} = 8.7$). Secondly, when we analysed only the trials in which participants were successful in the memory task and managed to judge 8 digits, the same repeated measure Bayesian ANOVA showed evidence against an effect of digit position ($BF_{\text{exclusion}} = 10.7$), against an effect of list length ($BF_{\text{exclusion}} = 43$), and against an effect of image type ($BF_{\text{exclusion}} = 10.9$).

In addition, we also examined whether the Subsequent-RTs differed between Experiments 2A and 2B. To this aim, we applied a 2 (experiment: 2A or 2B) x 2 (image type: real or non-real) x 3 (list length: 2 to 4) x 3 (digit position: 2 to 4) repeated measure Bayesian ANOVA. The best model from this analysis included only the simple effect of list length ($BF_{10} = 1.4 \times 10^3$), with evidence against an effect of digit position ($BF_{\text{exclusion}} = 13.2$), against an effect of image type ($BF_{\text{exclusion}} = 12.8$), ambiguous results against an effect of experiment ($BF_{\text{exclusion}} = 2.8$), and evidence for an effect of list length ($BF_{\text{inclusion}} = 1.2 \times 10^3$). All statistical interactions including digits' position had evidence against them (all $BF_{\text{exclusion}} > 9$). The effect of list length was due to the fact that Subsequent-RTs were faster with higher memory load, and higher memory load trials (memory load of 3 and 4) were only present in long list-length trials. Since the digit position effect in the first analysis was due to the way we pooled the Subsequent-RTs, RTs for digits in position 2 and onward from the same parity phase and from the same participant were pooled together for further analyses.

Chapter 4.2: The Influence of Attentional Refreshing on Face Recognition

Attentional refreshing is a process in working memory (working memory) that uses central attention to maintain mnemonic traces active in working memory and shield them from temporal decay and interference (for a review, see Camos, Johnson, Loaiza, Portrat, Souza & Vergauwe, 2018). It is deployed during maintenance intervals when central attention is not occupied by another attentionally-demanding task (Barrouillet & Camos, 2015, 2021).

There are at least two main ways to manipulate attentional refreshing opportunities. Firstly, refreshing opportunities can be reduced by increasing the cognitive load (CL) of a concurrent task; CL referring to the amount of time during which a task occupies central attention divided by the total processing time allotted to the task (Barrouillet et al., 2007). In working memory tasks in which item presentation is interspersed by a concurrent attentionally demanding task (i.e., complex-span task), higher CL tasks induce poorer recall performance (Barrouillet, Bernardin & Camos, 2004), because higher CL tasks occupies central attention for longer, and thus less time can be dedicated to refreshing, compared to lower CL tasks.

Secondly, evidence for refreshing can be gathered by manipulating the memory load in a working memory span task and measuring its impact on the reaction times to the concurrent task. Because attentional refreshing is thought to be deployed in a sequential manner (all representations refreshed one after the other, Lemaire, Pageot, Plancher, & Portrat, 2018; Portrat & Lemaire, 2014), higher memory load will yield longer RTs to the concurrent task; the refreshing of the memory items postponing the processing of the distractors in the concurrent task (Camos, Mora, Oftinger, Mariz-Eslig, Schneider & Vergauwe, 2019; Vergauwe, Camos & Barrouillet, 2014).

4.2.1 The present study

It has been suggested that refreshing can use long-term memory representations to maintain information in working memory via a process akin to a redintegration process (Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart, 1997). In other words, refreshing would use information held in long-term memory to reconstruct working memory traces (Barrouillet & Camos, 2015). In a series of two experiments, we investigated this hypothesis through examining the maintenance of the human faces material that is poorly studied in the working memory literature despite it being highly frequent for humans. To manipulate the possible reliance on long-term memory, we varied the familiarity of faces through the other race effect (ORE). The ORE refers to the fact that faces from participants own ethnicity are better recalled than faces from another ethnicity. It has been shown that this effect relies on how much a type of face is encountered in day-to-day life (Stahl, Wiese, & Schweinberger, 2008; Stelter & Degner, 2018; Valentine & Endo, 1992). Since learning in long-term memory relies massively on the number of times a memorandum is encountered (or a class of memoranda) in day-to-day life (Johnson & Miles, 2019; Szmalec et al. 2009), the ORE effect effectively relies on difference in long-term memory representation between different kind of faces.

In Experiment 1, face ethnicity of the memoranda was orthogonally manipulated to the CL induced by a concurrent parity task (Figure 1A). In Experiment 2, face ethnicity was orthogonally manipulated to the memory load in a Brown-Peterson paradigm, where we also used a parity task a concurrent task (Figure 1B). Raw data can be found on the OSF page of this chapter:
https://osf.io/n9zsg/?view_only=988c86ed75234f2796ab4580f535c6e7.

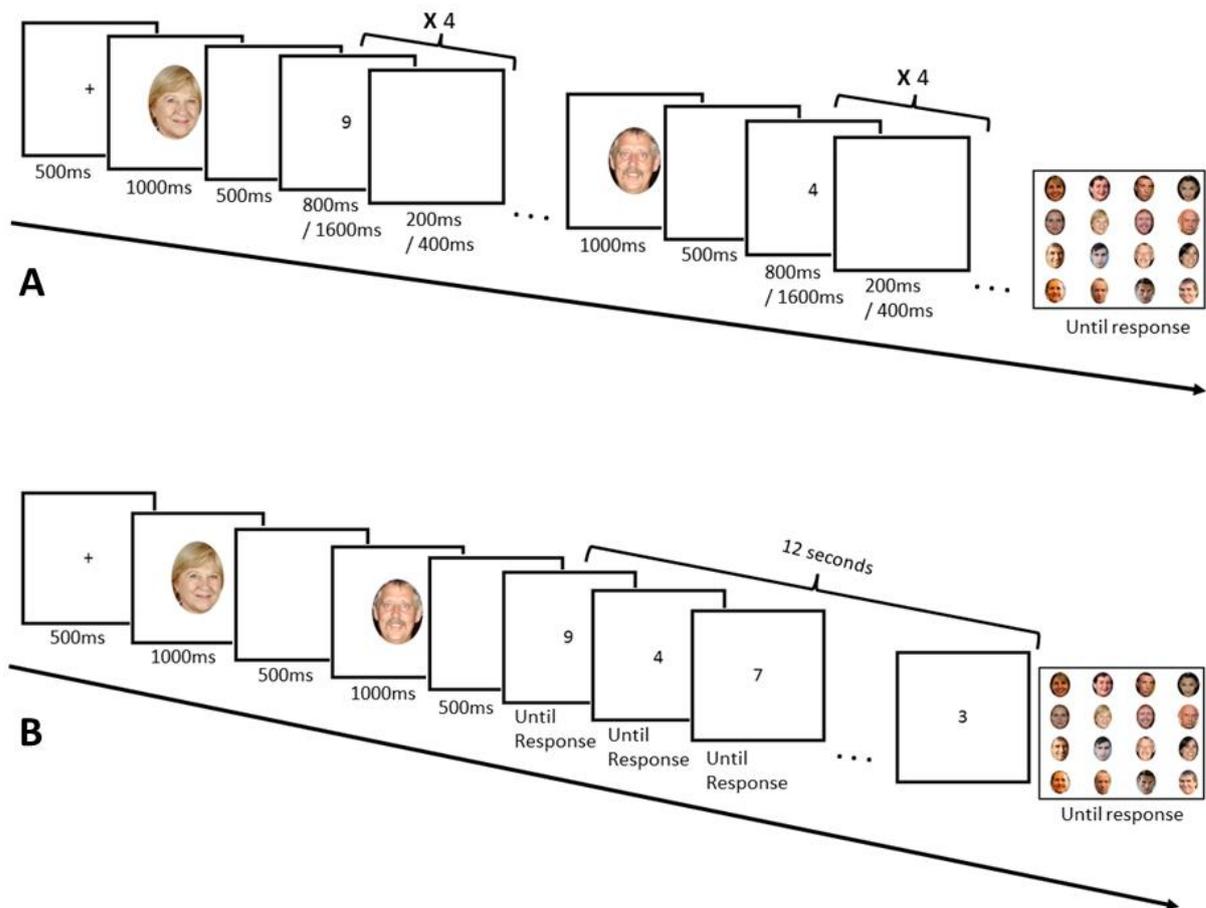


Figure 1: Schematic representation of a trial in experiment 1 (A) and experiment 2 (B). Both examples show trials with a memory load of two items and own-race, thus Caucasian used in example.

4.2.2 Experiment 1

In Experiment 1, we investigated how face familiarity influences recall performance in a complex span paradigm, in which the CL of the concurrent task was also manipulated. We firstly expect an effect of CL manipulation, with lower CL yielding better recognition performance compared to higher CL. Secondly, an effect of face familiarity on recognition performance would induce better performance for Caucasian faces compared to Asian faces. In addition, if familiar faces are refreshed more efficiently than less familiar faces, more familiar faces should benefit more from the refreshing opportunities in the low CL condition compared to the high CL condition. This would be marked by an interaction between the CL and the familiarity of the memoranda.

4.2.2.3 Methodology

Participants

Forty students from the University of Geneva (3 males, mean age: 20 ± 1.7 years) participated in this experiment. They were remunerated with partial course credits. Ethical approval for the study was given by IRB of the University of Geneva. Every participant to this study read and signed a form of consent. Since our participant sample was from Geneva (Switzerland), participants had to answer the following question: "have you lived for more than 3 months (cumulatively) in an East-Asian country (Korea, Japan, China, etc.)". Participants that answered yes to this question were discarded from the analysis. This was done to ensure that our participant's sample were not too familiar with Asian faces.

Material

The faces used as memoranda were selected from the 10k US Adults Faces database (Bainbridge, Isola, & Oliva, 2013). In this database, all faces were rated on their attractiveness, age and sex. Attractiveness and age were evaluated via a 5-points Likert scale. In addition, intrinsic memorability of each face was evaluated via a continuous recognition task. Series of faces were presented, and participants had to decide whether each face had already been presented or not. The rate of hit-rate and false alarm for each face was recorded. Sixty-three Asian faces were selected from the database. Each of them was paired with a Caucasian face that had similar scores on attractiveness, age, sex, and intrinsic memorability (as indexed by the ratio of hit-rate and false alarm, Table 1).

	Attractiveness	Age	Hit Rate	False Alarm
Asian Faces	2.89 ± 0.75	2.48 ± 0.80	0.61 ± 0.10	0.12 ± 0.08
Caucasian Faces	2.88 ± 0.81	2.50 ± 0.78	0.61 ± 0.10	0.12 ± 0.07

Table 1: Information on both pools of faces. Attractiveness: scale from 1 (not at all) to 5 (very much); Age: scale from 1 (very young) to 5 (very old); Hit rate: hit rate of correct recognition; False Alarm: Rate of incorrect recognition in a continuous recognition task.

Procedure

The experiment was divided in two blocks, one for each kind of faces. The block order was counterbalanced across participants. Each block used an increasing length procedure, where the number of images to maintain increased gradually from one face to six faces to maintain. There were six trials for each list length, three in each CL condition, totalling 36 trials per block. All blocks started with length-1 trials. The block stopped if no trial from a given length were successful or after participants finished trials with a list length of six items. The order in which high and low CL trials were presented was randomized for each list length.

We used a parity task as concurrent task. Four digits (1 to 9, 5 excluded) were presented sequentially on screen. Each digit stayed on screen for either 800ms followed by a 200ms ISI, or 1600ms followed by a 400ms ISI in the high and low cognitive load conditions, respectively. Digits were selected randomly. Participants were tasked to judge the parity of each digit. They responded by clicking the right mouse button for even digits and the left mouse button for the odd digits.

The experiment started with a training of the secondary task. Twenty-four digits appeared sequentially on screen (12 at the pace of each CL condition). Next, there were four training trials, two with 2 images to maintain and two with 3 images to maintain. Both Asian and Caucasian faces were used alternatively during the training trials. Faces used in the training were not used in the experimental trials.

Each trial started with a screen informing of the CL condition (High or Low) and participants had to press the space bar to start the trial. When they pressed the space bar, they had to start repeating "BA-BI-BU" out loud. Then, a fixation cross appeared in the centre of the screen for 500ms, followed by the first face presented for 1s. After the face disappeared and a 500ms blank screen, the concurrent task started. After the parity phase following the last presented image of the trial, participants were presented with an array of 16 faces and had to click on the presented faces in their order of presentation using the computer mouse. Participants

could not correct themselves and had to stop repeating “BA-BI-BU” as soon as the recognition phase started. After the participant had responded, the information screen for the next trial was prompted.

Within each block, each image was used twice as memorandum, once per CL condition. The images presented during response phase were intermixed with non-target images. Non-target images were randomly selected from the whole pool of images of the same type. Each image was used a maximum of 5 time as non-target and twice as memoranda.

4.2.2.4 Results

No participants were excluded on the basis of the post-experiments question. All participants reached at least 70% of correct responses to the parity judgement task, showing that they complied well with the instructions. Participants’ recognition performance was assessed via a span score. Each successful trial (i.e., all faces recognized in the correct position) was worth one point, and we then summed the points for each experimental cell and divided it by three (total number of trials per experimental cell). Span scores were analysed in a 2 (CL: high or low) X 2 (face type: Asian or Caucasian) repeated measure Bayesian ANOVA. The best model included only the main effect of face type, $BF_{10} = 2.8 \times 10^3$. There was extreme evidence for an effect of face ethnicity manipulation, $BF_{inclusion} = 2.8 \times 10^3$, with Caucasian faces yielding better performance than Asian faces (Mean span = 3.6 ± 1.1 and 3.1 ± 1.1 for Caucasian and Asian faces, respectively; Figure 2). However, evidence was gathered against an effect of CL on recall performance, $BF_{exclusion} = 5.8$ (Mean span = 3.3 ± 1.1 and 3.3 ± 1.2 for fast and slow pace, respectively). In addition, there were tentative evidence against the presence of the interaction, $BF_{exclusion} = 2.5$.

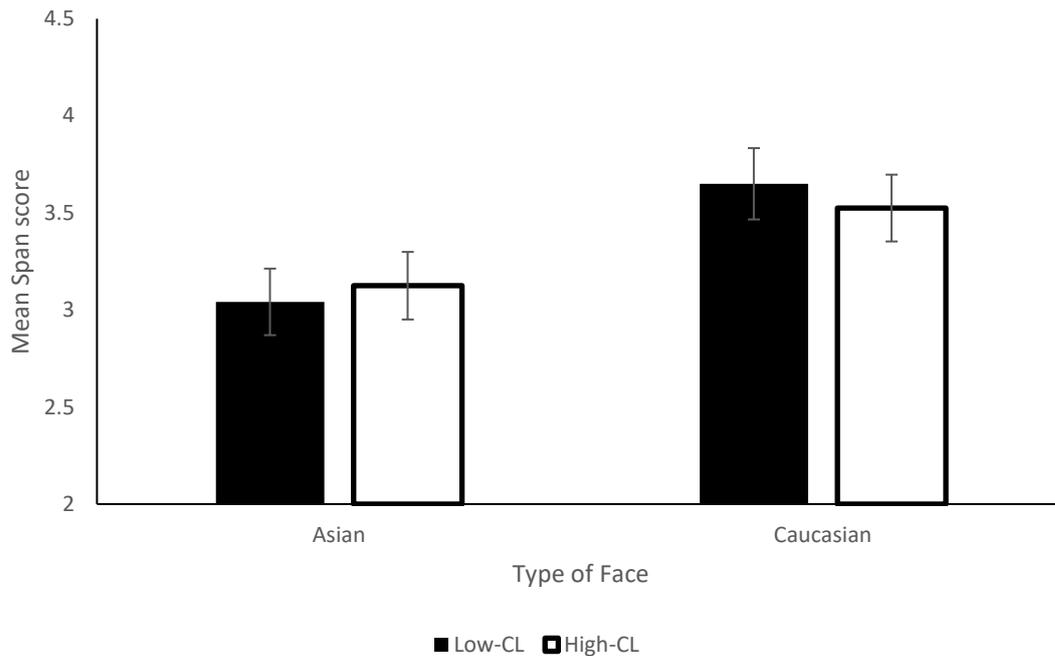


Figure 2: Mean span score as a function of concurrent task' cognitive load and type of face in experiment 1. Error bars are standard errors.

4.2.2.5 Discussion

In this experiment, we found better recall performance induced by Caucasian faces compared to Asian faces. Since we ensured that participants were more familiar with Caucasian faces than with east-Asian faces, our results are congruent with the presence of an ORE effect in our data. Interestingly, we found evidence against the concurrent task' CL manipulation. Although specific visuospatial material (unconventional visual characters, fonts of letters and position around a circle) have been shown to not be influenced by CL manipulations (Ricker & Cowan, 2010; Ricker & Vergauwe, 2020; Vergauwe, Camos & Barrouillet, 2014), most of the literature on attentional refreshing showed an effect of CL on recall performance. Finally, we found weak evidence against the interaction between CL manipulation and face familiarity.

The absence of CL effect on recall performance is a striking result. At least two main possibilities can explain this discrepancy with the literature. Either faces have a special status in working memory and cannot be refreshed, or the absence of effect is linked to specificities of our implementation of the cognitive load manipulation.

Indeed, it has been shown that faces have a special status in our cognitive system and may be processed differently than other kind of memoranda (Farah, Wilson, Drain, & Tanaka, 1998; Freiwald, Duchaine, & Yovel, 2016). To disentangle between the two possibilities, we designed a second experiment which investigated attentional refreshing without the use of CL manipulation.

4.2.3 Experiment 2

In Experiment 2, we implemented a Brown-Peterson task, in which the memory load as well as the familiarity of faces used as memoranda were manipulated. Brown-Peterson tasks are similar to complex-span tasks, with the difference that item presentation is grouped at the beginning of the trial and the concurrent task happens during a fixed duration following items presentation. As said in the introduction, manipulating memory load in a Brown-Peterson allow to investigate attentional refreshing. Since all memoranda will be refreshed between each processing episode, mean reaction time to the concurrent task will be postponed for each new memorandum to maintain. The hypothesis was that if more familiar faces are refreshed more efficiently than unfamiliar faces, then the memory load effect on reaction times to the concurrent task should be smaller for familiar compared to unfamiliar faces. The reason is that more familiar face would be refreshed more quickly, and thus postpone less the RT to the concurrent task. This would be marked by an interaction between memory load and face familiarity.

4.2.3.1 Methodology

Participants

Thirty-two students from the University of Geneva (8 men, mean age = 20.5 ± 2.2 years) participated in this experiment. They received partial course credits for their participation. None of them participated in the previous experiment.

Procedure

This experiment was divided in two blocks, one for each type of face. The block order was counterbalanced across participant. Each block followed the same increasing-length procedure as in Experiment 1, starting with a list length of one and with a maximum of six. There were six trials per list length. The block stopped if no trial for a given length was succeeded by the participant. In total, there was a maximum of 72 trials, 36 in each block.

Each trial started with a screen informing the participant of how many images would be presented in the current trial. Participants had to press the "space bar" to start the trial. After pressing the "space bar", they had to start repeating "BA-BI-BU" out loud and a fixation cross appeared in the middle of the screen for 500ms. Then the first face to memorize appeared on screen. Each face was presented for 1'000ms and was followed by a 500ms interval. After the last face was presented, the concurrent task started and was followed by the response phase. The response phase was the same as Experiment 1.

We used the same concurrent task as in Experiment 1, except that it lasted for a total of 12 seconds and participants were tasked to judge the parity of as much digits as possible during this interval. Each digit was replaced by the next one as soon as participants answered.

Before the experimental trials, participants had a training for the parity phase containing 20 digits that participants had to sort as odd or even. In addition, there were 4 training trials of the Brown-Peterson task, two with Caucasian faces and two with Asian faces as memoranda. As in Experiment 1, we asked participant whether they lived for more than three months in an East-Asian country. Participants that answered "yes" were excluded from analysis.

4.2.3.2 Results

The data of one participant was excluded because they answered “yes” to the post-experiment question. Of the remaining participants, none had less than 70% of correct responses to the parity judgement task. As in Vergauwe, et al. (2014), we also only kept the data of participants who had more than 50% of correct trials (all faces recalled in the correct position). The data of three participants were excluded based on this criterion, leaving us with a final sample of 28 participants.

We computed the span score for each participant as a function of type of faces using same calculation than in Experiment 1. We then used the span score as dependent variable in a paired Bayesian T-test with the type of faces (Asian or Caucasian) as independent variable (Figure 3).

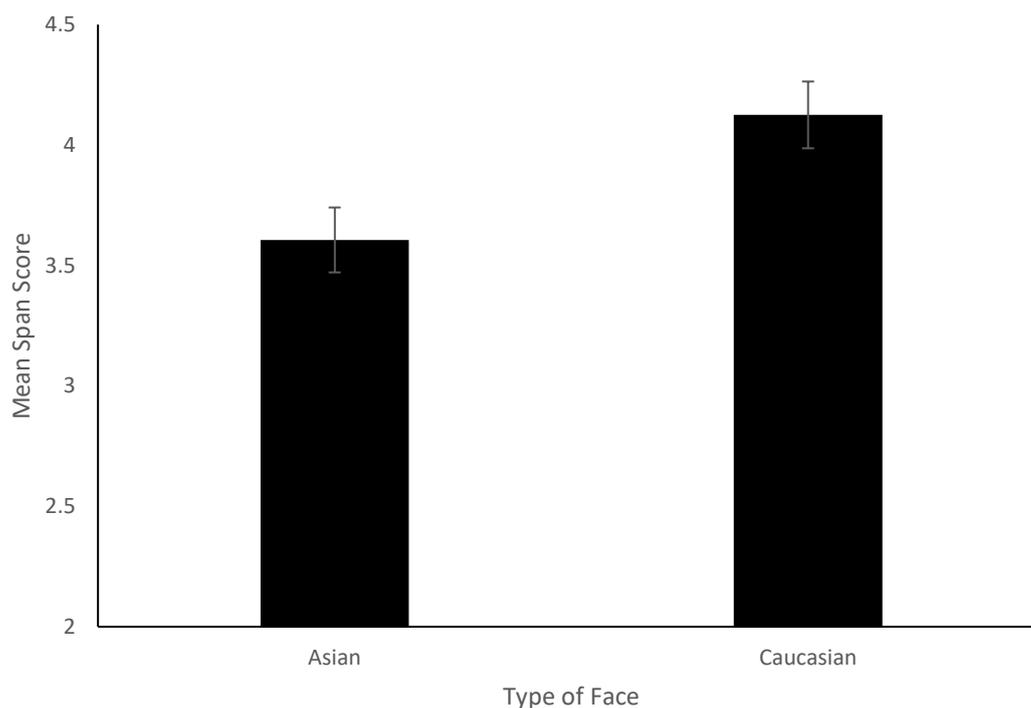


Figure 3: Mean Span score as a function of face type in experiment 2. Error bars are standard errors.

This analysis showed that Caucasian faces induced better recognition performance than Asian faces, $BF_{10} = 991$ (Mean span = 4.1 ± 0.7 and 3.6 ± 0.7 for Caucasian and Asian faces, respectively). We then analysed the reaction times to the parity task. As in Vergauwe et al. (2014), we separated the reaction times to the first

digit of the parity phase (hereafter: Initial-RTs) from the mean reaction times to the other digits (digits in position 2 and further, hereafter: Subsequent-RTs). Indeed, it has been shown that the Initial-RTs were always longer than all other RTs from the same processing phase, and thus should be analysed separately. In addition, we only kept memory load conditions in which more than 2/3 of participants succeeded to the memory task in 2 or more trials. Trials with memory load of 4 and further were discarded on this basis⁴. We then computed the mean Subsequent-RTs as a function of memory load and type of faces and used it as a dependent variable in a 2 (Type of face: Asian or Caucasian) X 3 (memory load: 1 to 3) repeated measure Bayesian ANOVA. This analysis showed evidence against both factors. There was evidence against the inclusion of the type of faces, $BF_{\text{exclusion}} = 6.1$ (Mean Subsequent-RTs Caucasian face: $651\text{ms} \pm 113\text{ms}$; Subsequent-RTs Asian face = $652\text{ms} \pm 107\text{ms}$), against an effect of memory load, $BF_{\text{exclusion}} = 5.0$ (Mean Subsequent-RTs 1 image: $663\text{ms} \pm 111\text{ms}$; Mean Subsequent-RTs 2 images: $652\text{ms} \pm 124\text{ms}$; Mean Subsequent-RTs 3 images: $642\text{ms} \pm 94\text{ms}$) and against the interaction, $BF_{\text{exclusion}} = 7.7$ (Figure 4).

⁴ Twenty-five participants succeeded in two or more trials with a list-length of 3; Fourteen succeed in 2 or more trials with list length of 4; Five participants succeeded in in or more trials with a list length of 5; 2 participants succeeded in in 2 or more trials with a list length of 6.

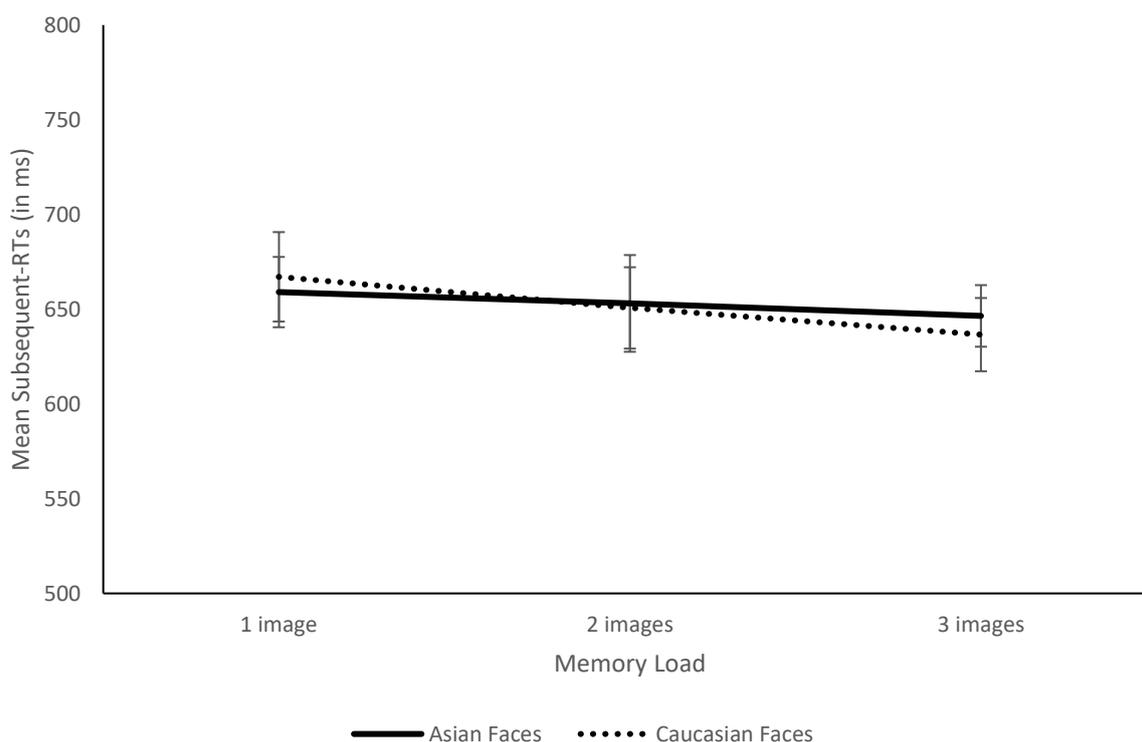


Figure 4: Mean Subsequent-RTs (in ms) as a function of memory load and type of face in experiment 2. Error bars are standard errors

Next, we analysed the Initial-RTs. As for the Subsequent-RTs, we only kept trials where participants succeeded in the memory task and kept only the memory load from 1 to 3. We then computed the Initial-RTs as a function of memory load and image type for all participants. Using this as a dependent variable in a 2 (type of face: Asian or Caucasian) X 3 (memory load: 1 to 3) Bayesian repeated measure ANOVA showed that Initial-RTs were impacted by the memory load only, $BF_{10} = 2.7 \times 10^4$ (Figure 5). We found evidence that smaller memory load induced slower Initial-RTs, compared to higher memory load, $BF_{inclusion} = 3.3 \times 10^4$ (Mean Initial-RTs 1 image: 1'027ms \pm 357ms; Mean Initial-RTs 2 images: 920ms \pm 299ms; Mean Initial-RTs 3 images: 812ms \pm 239ms). We found weak evidence against an effect of type of faces, $BF_{exclusion} = 2.5$ (Mean Initial-RTs Asian face: 964ms \pm 346ms; Mean Initial-RTs Caucasian face: 913ms \pm 316ms), and substantial evidence against the interaction, $BF_{exclusion} = 6.0$.

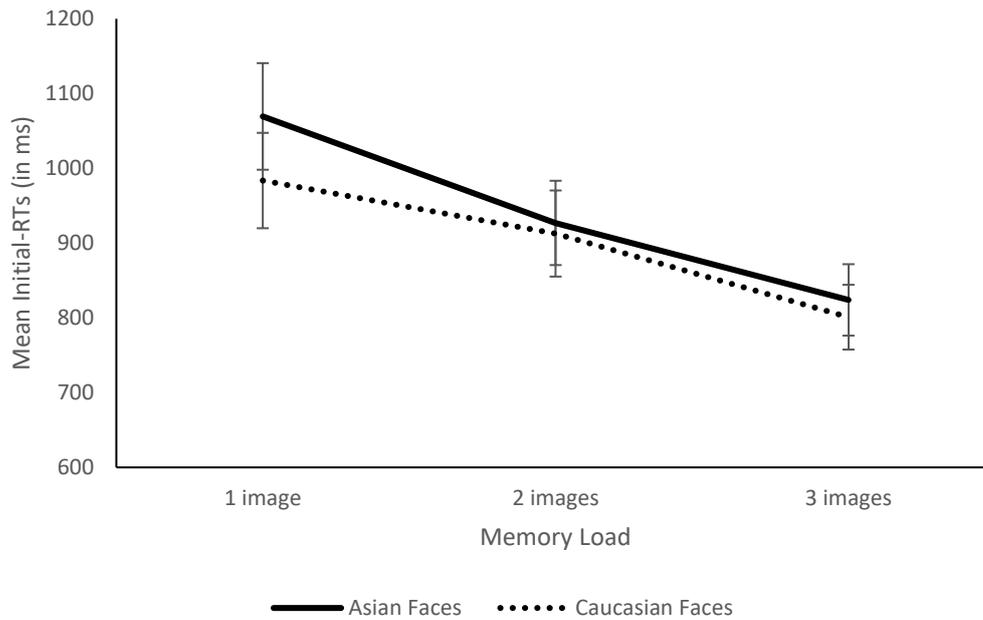


Figure 5: Mean Initial-RTs (in ms) as a function of memory load and type of face in experiment 2. Error bars are standard errors.

4.2.3.3 Discussion

In Experiment 2, analysis of the span score replicated the ORE effect, whereby faces from participant’s ethnicity were better recalled than faces from another ethnicity. Regarding the analysis of the reaction time, we found no effect of memory load or type of face manipulation on the mean Subsequent-RTs. Since the effect of memory load on concurrent task reaction time has been used as an index of attentional refreshing functioning (Vergauwe, Camos & Barrouillet, 2014; Camos et al., 2019), our data goes against the hypothesis that more familiar faces are refreshed more efficiently than less familiar faces. More precisely, results from Experiment 2 are congruent with the hypothesis that faces are not maintained via attentional refreshing at all. Finally, analysis of the Initial-RTs showed that this reaction time was influenced by the memory load, where higher memory load induced shorter Initial-RTs. This is a striking result, as other studies showed the opposite effect: an increase in Initial-RTs linked to an increase in memory load. It may be that the use of an increasing length procedure induced this pattern of results.

Since trials with higher memory load were always at the end of a block, participants had time to improve their performance to the concurrent task, compared to trials with lower memory load which were always at the beginning of a block. Participants could then have more training on the concurrent task during higher memory load conditions compared to lower memory load conditions. This would mean that they would answer faster to the concurrent task in later trials, which correspond to higher memory load trials. If participants were indeed faster at the concurrent task the further, they were in a block, then mean Subsequent-RTs should also be shorter at the end of a block compared to the beginning of a block. This effect would go in the opposite direction than the expected effect of memory load on Subsequent-RTs, which predict an increase in reaction time induced by higher memory load. Thus, it may be that faces were effectively refreshed during the maintenance interval, but the impact of attentional refreshing on Subsequent-RTs was masked by the training effect induced by the increasing-length procedure.

4.2.4 General discussion

In two experiments, we replicated the ORE. Faces from participant's ethnicity were better recalled than faces from another ethnicity. Interestingly, we gathered evidence against an effect of attentional refreshing opportunity manipulation on face recognition. There was evidence against an effect of CL manipulation on recall performance in experiment 1, and against an effect of memory load manipulation on reaction time to the concurrent task in experiment 2; Both expected effects being indexes of attentional refreshing functioning. Taken together, these patterns of results are congruent with the idea that faces, in general, are not maintained via attentional refreshing.

However, as indicated above, this conclusion may be limited due to the actual implementation of experiment 2. One way to disentangle the possible effect of memory load and training on the concurrent task would be to not use an increasing-

length procedure. Thus, the possible training effect on the concurrent task would not be confounded with memory load manipulation.

Nonetheless, irrespective of the possible implication of a training effect in our reaction time measurements in Experiment 2, the absence of a CL effect in Experiment 1 is still striking. In the attentional refreshing literature, most material have been shown to be sensitive to CL manipulation. Materials that seem to not be influenced by CL manipulation were always mainly uncommon memoranda (fonts, unconventional characters, position around a circle), and always in the visual domain (Ricker & Cowan, 2010; Ricker & Vergauwe, 2020; Vergauwe, Camos & Barrouillet, 2014). This has led researchers to postulate that unrefreshable memoranda had little or no prior representation in long-term memory (Ricker & Vergauwe, 2020). Regarding the type of memoranda we used in the current study, faces cannot be considered as uncommon, as they are ubiquitous in our day-to-day lives. It is then difficult to explain that faces are not refreshable due to their uncommonness. This gives credence to the hypothesis that faces have a special status in our cognitive system.

One other possible explanation to our pattern of results relies on the possibility that, for very familiar memoranda, active maintenance (like attentional refreshing) can be replaced by simply relying on long-term memory. A recent study showed evidence that long-term memory representation could replace active maintenance in condition where the memoranda does not suffer from interference and has a stable long-term memory representation (Schurgin, Cunningham, Egeth & Brady, 2018). It may be that faces were not actively maintained via attentional refreshing, because faces are a very common type of memoranda. In addition, since we used a recognition paradigm, participants could easily rely, at least partly, on a feeling of familiarity to recognize the presented faces, which use other cognitive mechanisms than direct recollection (Migo et al., 2009). This would make the memory task easier and thus could mask the effect of CL manipulation in Experiment 1.

4.2.5 Conclusion

Our results leave two main possibilities regarding the relationship between attentional refreshing and the maintenance of faces. Either faces are refreshed but our experimental implementations did not manage to detect it, or faces are indeed unrefreshable. Regarding the first possibility, the same methodology used in Experiment 1 has reliably shown an attentional refreshing opportunity effect on recall and recognition performance with verbal as well as visuospatial material (Vergauwe et al. 2014). That would mean that faces are refreshed much more quickly than other types of material, because our implementation did not manage to detect the effect. It would also mean that the effect of memory load on Subsequent-RTs was masked by the training effect induced by the increasing length procedure in Experiment 2.

If, on the contrary, faces simply cannot be refreshed, it would not be congruent with the hypothesis that only memoranda with poor long-term memory representation cannot be refreshed. Indeed, faces are very common in day-to-day lives. In addition, there was evidence against an effect of attentional refreshing opportunity manipulation on both familiar and unfamiliar faces. Future experiments should aim at generalising our results, by using other faces and concurrent task. Regarding the second possibility, if faces indeed cannot be refreshed, this would mean that faces, as memoranda, have a special status in our cognitive system.

Chapter 5: Does Visuospatial Familiarity Influence Consolidation Speed?

Working memory is one of the most studied constructs in the field of cognitive psychology. Indeed, its role as the center of online processing and maintenance of information has been shown to be predictive of many higher-level cognitive functions (Cowan, et al., 2005; Kail, 2007; Unsworth & Engle, 2007). Even though many aspects of its inner functioning are still a matter of debate, modern models of working memory all acknowledge that attention plays an important role, especially during maintenance. More precisely, several processes have been proposed to explain how availability of attentional resources influences working memory functioning. In this chapter, we will specifically investigate one of these processes, the short-term consolidation. Short-term consolidation (or consolidation in working memory) is thought to transform fleeting iconic memory traces into more stable working memory representations (Jolicoeur & Dell'Acqua, 1998; Ricker & Cowan, 2014; see Ricker, Nieuwenstein, Bayliss & Barrouillet, 2018, for a review). Although several studies investigated this construct and how it operates, its inner functioning is still debated. One possibility about consolidation is that it relies, at least in part, on semantic long-term memory in a process akin to a redintegration process (Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart, 1997). This supposes that information that is already well defined in semantic long-term memory would be consolidated more efficiently than information that is less well-represented in semantic long-term memory. The aim of the present study was to directly investigate this possibility by varying consolidation opportunity orthogonally to the familiarity of visual memoranda.

5.1 Consolidation in working memory

Consolidation in working memory was first proposed by Jolicoeur and Dell'Acqua (1998). In their experiments, a set of memoranda (letters or symbols)

appeared on screen for a fixed time and was followed by an auditory signal to which participants had to respond to as fast as possible. The authors manipulated the stimulus onset asynchrony (SOA) between memoranda presentation and the auditory signal, as well as the number of presented items during a trial. In addition, they also manipulated the instructions to either memorize the stimuli or pay no attention to them. The results showed that, when participants had to memorize the presented items, shorter SOA induced longer reaction times to the auditory cues. Interestingly, there was no effect of the SOA manipulation when participants did not have to memorize the presented items. The set size manipulation also had an impact on the reaction times to the auditory signal; reaction times were significantly longer when the set size was larger.

Since the effect of SOA manipulation on reaction time to the auditory cue only happened when participants were tasked to remember the identity of the letter and it was also altered by the set size, the authors concluded that a consolidation process was happening during the free time in between item presentation and the auditory cue. Considering the results, the authors also surmised that this process demanded central attention and was time consuming, otherwise no effect of SOA manipulation should have been observed on the reaction times to the auditory signal. Additionally, the effect of set size on the reaction times could be explained by supposing that this process had a capacity limit (although the exact capacity was not evaluated directly in their experiments).

Follow-up experiments by Stevanovski and Jolicoeur (2007), using the same basic paradigm as Jolicoeur and Dell'Acqua (1998), showed that blocking articulatory rehearsal by the use of articulatory suppression did not make the effect of the mask SOA manipulation on reaction times disappear. This was congruent with the idea that the consolidation process likely relied on central attention and was independent of rehearsal.

However, these experiments did not show that the process responsible for the effect on reaction times was directly involved in memory creation. Indeed, the experiments presented above showed an effect of mask SOA manipulation only on the reaction times to the concurrent task. If the interval between item presentation and the mask are indeed used to consolidate information in working memory, then disturbing the consolidation process should also impact recall performance. Nieuwenstein and Wyble (2014) tested this hypothesis by using the same dual task methodology as Jolicoeur and Dell'Acqua (1998) but with more difficult memory and concurrent processing tasks. In their experiment, memoranda comprised of either four letters or one unconventional complex visual shape (Chinese characters). They also manipulated the presence or absence of a mask following image presentation and the SOA at which the mask appeared on screen. They used a parity task as concurrent task. It has been shown that a parity task taps into central attention (Barrouillet et al., 2007) and thus can be considered as more difficult than the concurrent task used by Jolicoeur and Dell'Acqua (1998). Their results showed that recall performance was impacted by the delay manipulation: the shorter the SOA between item presentation and a mask, the worse the recall performance to the memory task. In addition, shorter SOA between item presentation and the mask induced longer RT to the concurrent task.

The impact of the delay manipulation on recall performance was also shown in another study. In a series of experiments, Bayliss, Bogdanov and Jarrold (2015) investigated whether consolidation and attentional refreshing were separate processes. Attentional refreshing is another process in working memory that also uses central attention to maintain mnemonic traces active in working memory (see Camos et al. 2019 for a review). Since attentional refreshing happens between processing episode, it can be manipulated by varying the cognitive load of a concurrent task (the amount of time a concurrent task occupies central attention). To this aim, the authors used a complex-span task (i.e., a span task in which memoranda are presented sequentially, each of them followed by a concurrent processing episode) in which

they manipulated when the concurrent task started. The concurrent processing task could happen directly after item presentation or after an unfilled interval. The concurrent task could start either 200 milliseconds after item presentation or after a delay of 2'400 milliseconds (i.e., consolidation interval). In addition, they also manipulated the difficulty of the concurrent task. Participants had to either read additions aloud with the results presented on screen (low difficulty condition) or read the same additions aloud and calculate the result (high difficulty condition). The results showed that recall performance was worse when the interval after item presentation was short, compared to when it was long. The authors also showed that a more difficult concurrent task induced worse recall performance than a less difficult concurrent task. However, they found no interaction between both factors. They thus concluded that consolidation and attentional refreshing were independent processes, although they tap on the same resource.

5.2 Consolidation and encoding processes

Although the studies presented above have been seminal for proposing the construct of consolidation in working memory, their operationalization of the consolidation opportunity is open to debate. In these experiments, the consolidation manipulation consists of manipulating the delay between item presentation and the beginning of a concurrent task or the appearance of a mask. However, encoding processes also take place during the first few hundreds of milliseconds following image presentation (Bays et al., 2011). Encoding is the set of processes used to primarily construct elements of our environment into a mental construct that can then be manipulated by our cognitive system. Manipulating free time directly after memoranda presentation makes it difficult to disentangle if the manipulation affects encoding processes, consolidation, or both. Thus, conclusions from the presented-above studies could concern encoding processes as much as consolidation process.

Nonetheless, several studies using different approaches have shown that encoding processes and consolidation are separable. In two separate studies, Chen

and Wyble (2015, 2016) asked participants to memorize the location of a letter in an array of numbers. A mask followed memoranda presentation after a fixed delay. Then, participants had to recall either the identity of the letter, the location of it or both information. The interesting manipulation is that, in some trials, participants were not instructed to memorize the identity of the letter but were asked for it at the end of the trial anyway. Results showed that when participants were tasked to memorize the identity of the letter, they could recall it without any problem. In contrast, in trials where they were not asked to memorize the identity, participants could recall its location but rarely its identity. Since the letter had to be processed at some level to determine if it was a letter or a number, it had to be encoded. But since participants could not recall its identity when not asked for it, the authors argued that this information was not consolidated, even though it was encoded. This pattern of results is congruent with the supposition that encoding and consolidation processes are separated. It also shows that consolidation is goal-directed, although maybe not directly consciously controlled.

Other studies point in the direction of a separation between consolidation and encoding processes. Several studies operationalized the consolidation delay as the delay between a mask directly following item presentation and the beginning of a concurrent task or the appearance of a probe (hereafter: post-mask delay manipulation, in contrast to pre-mask delay manipulation, Figure 1). The idea is that a mask would overwrite the iconic traces induced by item presentation. Since encoding

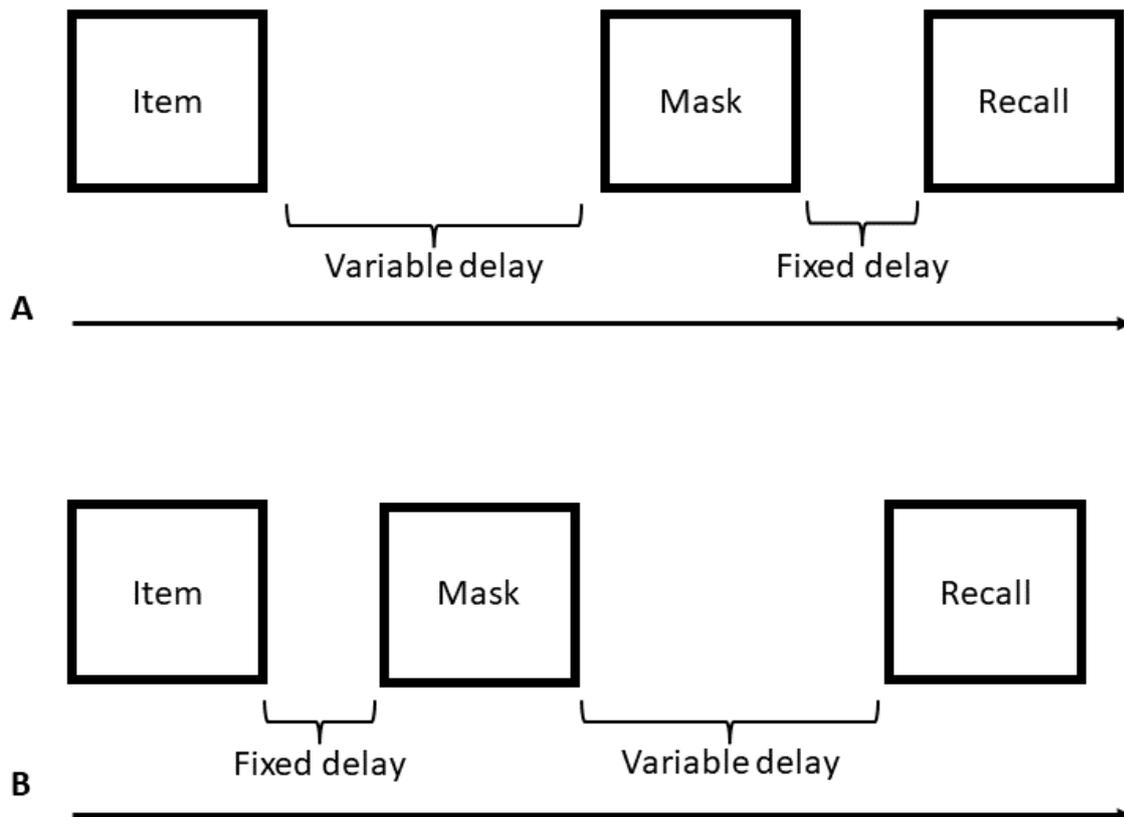


Figure 1: Schematic representation of a pre-mask manipulation (A) and a post-mask manipulation (B) process relies on iconic memory (Tripathy & Ögmen, 2018), masking would stop encoding of the presented memoranda. Manipulating a post-mask delay would then allow to manipulate only consolidation and not possible encoding processes. Following this logic, De Schrijver and Barrouillet (2017) found a beneficial effect of post-mask delay manipulation on recall performance in a complex-span task. In their task, the memoranda were series of letters, and the concurrent task was a parity judgment task (i.e., to decide whether a presented number was odd or even). Their aim was similar to the Bayliss et al. (2015) study. They investigated whether consolidation and attentional refreshing were two separable processes or not. The

main difference with the study from Bayliss et al. (2015) is that De Schrijver & Barrouillet (2017) manipulated consolidation via a post-mask manipulation, instead of a pre-mask delay manipulation in the Bayliss et al. (2015) study. De Schrijver & Barrouillet (2017) also manipulated the difficulty of the concurrent task orthogonally to the post-mask delay. The consolidation interval was operationalized as the time between a mask directly following memoranda presentation and the start of the concurrent task.

There were two main differences with the Bayliss et al. (2015) experiments. Firstly, in the De Schrijver & Barrouillet (2017) study, item presentation was directly followed by a mask. Secondly, the consolidation interval was manipulated with more than two levels. The concurrent task could start 0, 500, 1'000, 2'000, 3'000 or 5'000 milliseconds after the mask disappeared from screen, whereas, in Bayliss et al. (2015), the concurrent task started either 200 or 2'400 millisecond after the mask disappeared from the screen. The results from De Schrijver & Barrouillet (2017) show that recall performance is influenced by the consolidation interval manipulation, with a shorter consolidation interval leading to worse recall performance, compared to a longer consolidation interval. In addition, recall performance was also influenced by the difficulty of the concurrent task, with lower recall performance induced by more difficult concurrent task. Regarding their initial hypothesis, the authors only found an interaction between consolidation and attentional refreshing manipulation when participants were not under articulatory suppression. When participants were under articulatory suppression, they found no evidence for the interaction between attentional refreshing and consolidation manipulation. The authors thus concluded that consolidation and refreshing were separable processes because the interaction was only evident when verbal maintenance processes could be used. The results show that a longer post-mask interval was beneficial for recall performance. This benefit is separated from encoding process, since it happens after a mask, and the mask stops encoding of the memoranda.

Regarding the relationship between encoding and consolidation processes, one last line of research argues in favor of their separation. It has been shown that consolidation process is not directly influenced by physical properties of the memoranda. In a change detection task, Sun, Zimmer and Fu (2011) orthogonally manipulated the SOA between a mask and a probe with the visual complexity of Chinese characters in two different participant groups. Participants could either be novices (Germans) or experts (Chinese reader) in Chinese language. Participants were presented with an array of three Chinese character, then a mask appeared on screen after a variable delay (their SOA manipulation) and finally a probe containing three Chinese characters appeared on screen for two seconds. Participants had to determine whether the test array was the same as the presented one at the beginning of the trial. They found that expert Chinese readers had overall better recall performance than novice Chinese readers. They also found that longer SOA delay induced better recall performance than shorter SOA. Interestingly, their results showed that visual complexity and SOA manipulation did not interact together. It has been shown that physical properties of visual memoranda do impact encoding processes, with more complex memoranda needing more time or resources to be encoded than less complex memoranda (Eng, Chen & Jiang, 2005; Kursawe & Zimmer, 2015). The absence of an effect of physical complexity on consolidation is congruent with the supposition that both systems are separable. Indeed, if both systems were not separable, they should be influenced by the same factors.

Taken together, the studies presented above put in light several aspects of consolidation in working memory. First, consolidation seems separated from encoding processes. Secondly, consolidation imposes a processing bottleneck in working memory. This means that consolidation postpones the processing of other information while it is used. If this was not the case, we would not detect an effect of consolidation interval on the RTs of a concurrent task. In addition, the beneficial effect of a longer interval between item presentation on recall performance shows that consolidation is a process that helps creating a working memory representation.

Lastly, the combination of the bottleneck effect and the fact that the effect of consolidation interval is not influenced by the presence or absence of articulatory suppression shows that consolidation relies on a central attentional resource. This is also corroborated by the seemingly domain-general aspect of consolidation, since the effect of consolidation interval has been shown with letters (De Schrijver & Barrouillet, 2017, Nieuwenstein & Wyble, 2014) as well as with visuospatial memoranda (Ricker & Cowan, 2014; Sun, Zimmer & Fu, 2011).

5.3 Is consolidation influenced by familiarity?

Although some aspects of consolidation are known, there are still many unknowns on its general functioning. As we saw in the preceding sections, consolidation seems to be domain-general, as consolidation delay manipulation impacts verbal as well as visuospatial material. Nonetheless, it might be that some classes of memoranda are more difficult to consolidate than other classes of memoranda. Regarding this possibility, it has been postulated that more familiar memoranda would be consolidated more quickly and efficiently than less familiar memoranda (Blalock, 2015; Xie & Zhang, 2017). Two studies aimed at investigating this possibility. However, they both manipulated a pre-mask delay and not a post-mask delay. Thus, both experiments confounded consolidation and encoding processes.

The first study examined was done by Xie and Zhang (2017). In a series of experiments, they manipulated the delay between a memoranda presentation and the appearance of a mask on screen in a change detection task. The mask could appear either 117 or 340 milliseconds after the memoranda presentation and the total interval between memoranda presentation and the probe appearance was 1'600 milliseconds. In addition, there was a "no-mask" condition, in which the 1'600-millisecond interval between memoranda presentation and the probe remained unfilled. To manipulate familiarity, two groups of participants underwent the

experiments. One group of participants was familiar with the memoranda whereas the other group was not familiar with it.

As memoranda, the authors used image of monsters coming from the Pokémon game. Participants who were already familiarized with Pokémon were assigned to the "familiar" group, whereas participants that did not know them well were assigned to the "non-familiar" group. Since new generations of Pokémon games are coming out approximately every two years, the authors also manipulated the familiarity of the memoranda itself with two levels, by manipulating the Pokémon generations. These could either come from the first generation of Pokémon (with whom the "familiar" group was well acquainted) or coming from the latest generation at the time (with whom the "familiar" group was less acquainted). The "non-familiar" group was as well acquainted with Pokémon coming from both generations. The results showed that shorter delay between memoranda presentation and the mask appearance was linked to poorer recognition performance in both groups, replicating the effect of pre-mask SOA manipulation found in other studies. In the "non-familiar" group, there was no effect of which generation Pokémon were coming from, both group of Pokémon suffered identically from shorter SOA. This was in contrast to the "familiar" participant group, in which the manipulation of the Pokémon's generation and the mask delay manipulation significantly interacted. In the longer delay condition, Pokémon coming from the first generation induced significantly better recall performance compared to Pokémon coming from the latest generation. This difference in recall performance disappeared in the condition with the shorter delay between item presentation and the mask. The difference between Pokémon generation on recall performance was present only in the group that had a differing long-term memory representation about Pokémon generation (i.e., familiar group). That was not the case in the group that was not familiar with both groups of Pokémon (i.e., the non-familiar group). The authors concluded that more familiar memoranda were consolidated more efficiently than less familiar memoranda, and that was one of the factors underlying the familiarity advantage in working memory.

The second study investigating visual familiarity and consolidation was done by Blalock (2015). Contrary to Xie and Zhang (2015), Blalock (2015) manipulated visual familiarity by training participants on specific shapes and then manipulating whether trained or untrained shapes were used as memoranda in a change detection task. Consolidation was manipulated by varying the delay between item presentation and the appearance of a mask. The results showed that accuracy was higher for trained shapes compared to untrained shapes and that unfamiliar shapes suffered more from shorter SOA compared to familiar shapes. This pattern of results was congruent with the results from Xie and Zhang (2015). More familiar items were better recalled than less familiar items, and the difference was more pronounced with longer SOA compared to shorter SOA. This interaction is explained by authors of both articles as evidence for an impact of visual familiarity on consolidation efficiency. More familiar memoranda are consolidated more efficiently than less familiar memoranda, and thus they are better consolidated after longer SOA than less familiar memoranda. In shorter SOA conditions, familiar memoranda can benefit less from their increased consolidation, and thus induce less difference on recall performance compared to less familiar memoranda.

However, as said earlier, both studies presented above manipulated consolidation by varying a pre-mask delay. As discussed above (chapter 5.2), manipulating consolidation this way makes it difficult to disentangle whether the delay manipulation impacted consolidation in itself or encoding processes as well. In addition, Blalock (2015) manipulated the impact of long-term memory by training the participants just before they were tested on these shapes, and they used a change detection task. In a change detection task, participants do not have to completely recall the probed memoranda, as they can answer by using a visual pattern-matching process instead of directly giving back the whole memorized item. Thus, the training used in this experiment could be not effective enough to really create a stable representation of the trained shape in long-term memory, but effective enough to

create a feeling of familiarity with these shapes, which could be sufficient to perform in the task.

Furthermore, the Xie and Zhang (2015) and Blalock (2015) studies concluded that more familiar memoranda were consolidated more quickly than less familiar memoranda by evaluating the impact of the SOA manipulation on recall performance. This can be considered as indirect evidence of a change in consolidation speed, but not as a direct evidence since reaction times were not directly measured. Their results can also be explained if we consider that more familiar memoranda are consolidated at a deeper level than less familiar memoranda. It could be that the consolidation process enriches the consolidated information by using already existing information in long-term memory, similarly to a redintegration process. A redintegration process is a process via which representation in working memory is reconstructed by using information held in long-term memory (Hulme, et al. 1997). If this is the case, more familiar memoranda would be consolidated to a higher level than less familiar memoranda, but not more quickly.

5.4 The present study

In the present study, we investigated whether visual familiarity impacts consolidation in working memory. More precisely, we investigated whether more familiar memoranda are consolidated more quickly than less familiar memoranda. In a series of two experiments, we used a complex span task in which we manipulated the delay between a mask following memoranda presentation and the beginning of a concurrent attention-demanding task (here after: consolidation delay manipulation). The idea was that, if consolidation is an attention-demanding task and that processing imposes a processing bottleneck in working memory, then reaction times to the concurrent task should be affected by the consolidation manipulation. Shorter consolidation delay will induce longer reaction times to the concurrent task because processing of the concurrent task will be postponed until consolidation process has been sufficiently completed. This logic was used in several studies that aimed at

investigating the impact of a memory load on reaction times to a concurrent processing task (Camos, Mora, Oftinger, Mariz-Elsig, Schneider & Vergauwe, 2019; Schneider, Vergauwe & Camos, in prep; Vergauwe, Camos & Barrouillet, 2014).

These studies showed that, in tasks in which memoranda presentation is directly followed by a concurrent processing task, the reaction times to the first processing episode (hereafter: Initial-RT) is longer than the following reaction times from the same processing episode: Therefore, these processes should be analyzed separately from the other processing episodes. In conditions with shorter consolidation delay, consolidation process should not be over when the concurrent task starts, and thus processing of the concurrent task will be postponed until consolidation is over. In contrast, in conditions with longer consolidation delay, the memoranda should be already consolidated when the concurrent task starts, and thus the reaction times to the first digit should be quicker than in trials with shorter consolidation delay.

In addition to the consolidation manipulation, the familiarity of the memoranda was also directly manipulated. In Experiment 1, we manipulated visual familiarity in a within-subject design, by varying the degree of familiarity between two sets of images. Regarding recall performance, our hypothesis was that high familiarity images should induce better recall performance than low familiarity image, and shorter consolidation delay should induce worse recall performance than longer consolidation delay. If consolidation process is influenced by long-term memory representation, then we should also detect an interaction. The difference between low and high familiarity images should be bigger with longer consolidation delay, compared to shorter consolidation delay. This would mean that more familiar images can benefit more from consolidation delay than less familiar images.

We also hypothesized that Initial-RTs to the concurrent task should be influenced by familiarity and consolidation delay manipulation. If more familiar images are consolidated more quickly, then Initial-RTs should be quicker for more

familiar images compared to less familiar ones, as consolidation process should be over more quickly for high familiarity image compared to less familiar images. In addition, Initial-RTs should be longer when the consolidation delay is shorter.

In Experiment 2, the familiarity was manipulated between-subjects. Half of the participants underwent a series of training sessions in which they had to learn some abstract figures (i.e., Trained group), whereas the other half of the participants did the experimental tasks without prior knowledge about the abstract figure they had to memorize (i.e., No-training group). This way, we could directly manipulate difference in long-term memory representation between both group of participants. Our hypotheses for this experiment were similar to those in Experiment 1. The trained group should have better recall performance than the no-training group, and shorter consolidation delay should induce worse recall performance than longer consolidation delay. If consolidation is influenced by item familiarity, then the difference between groups should be bigger in condition with longer consolidation delay, and this would be marked by a statistical interaction between both factors.

Regarding reaction times, shorter consolidation delay should induce slower Initial-RTs compared to longer consolidation delay, and the trained group Initial-RTs should be lower than the no-training group. An interaction between both factors, where the difference in Initial-RTs is bigger in longer consolidation delay, compared to lower consolidation delay, would be congruent with the hypothesis that more familiar images are consolidated more quickly than less familiar images. The idea is that, if more familiar memoranda are consolidated more quickly than less familiar memoranda, then Initial-RTs should be slower for less familiar memoranda, compared to more familiar images. This would mean that consolidation process would be over more quickly for more familiar memoranda, compared to less familiar ones.

Since we wanted to gather evidence for or against the presence of an effect, this study uses the Bayesian statistical approach. In the Bayesian approach, the resulting statistic is a number comprised between 0 and positive infinity. This number

can be interpreted straightforwardly as how many times a model explains the data better (or worse) than another model. For example, a model with a Bayes factor of 10 when compared to the null model implies that this model explains the data 10 times better than the null model. The Bayes factor can be calculated at the model level or at the variable level. When calculated at the variable level, it is called the Bayes factor for the inclusion (or exclusion) of a factor. This is the Bayes factor of the model with the variable of interest divided by the Bayes factor of the model without this variable of interest. The Bayes factor for (or against) the inclusion of a variable can also be interpreted straightforwardly as how much evidence there is in the data to include (or exclude) the variable. All analyses in this study were performed with the BayesFactor package (Morey & Rouder, 2018) in R (R Core Team, 2020), with default settings.

5.5 Experiment 1

In this experiment, we used a complex span task in which we manipulated the SOA between a mask directly following image presentation and the beginning of a concurrent processing task (hereafter: consolidation delay manipulation, 3 levels: 100ms, 500ms and 1'000ms). Orthogonally to the consolidation delay manipulation, we also manipulated the familiarity of the images to memorize (2 levels: Familiar or Unfamiliar). To minimize the use of verbal rehearsal, participants were under articulatory suppression (repeating "BA-BI-BU" out loud) during item presentation and the processing task.

5.5.1 Method

Participants

Thirty-six students from the University of Fribourg (32 women, mean age: 20.7 ± 1.8 years, range: 18-27) participated in this experiment. They were remunerated with partial course credits or cinema tickets. Every participant read and signed a form of consent. Ethical approval for this study was given by IRB of the University of Fribourg.

Material

The stimuli for the familiar condition consisted of 42 objects taken from Snodgrass and Vanderwart (1980) image database (Figure 2). We selected images related to concepts with high lexical frequency nouns to ensure that the images represented well-known objects. The stimuli for the unfamiliar condition comprised 42 images that were created from smoothly connected features of the Snodgrass and Vanderwart image set (taken with permission from Soldan, Zarahn, Hilton & Stern, 2008).

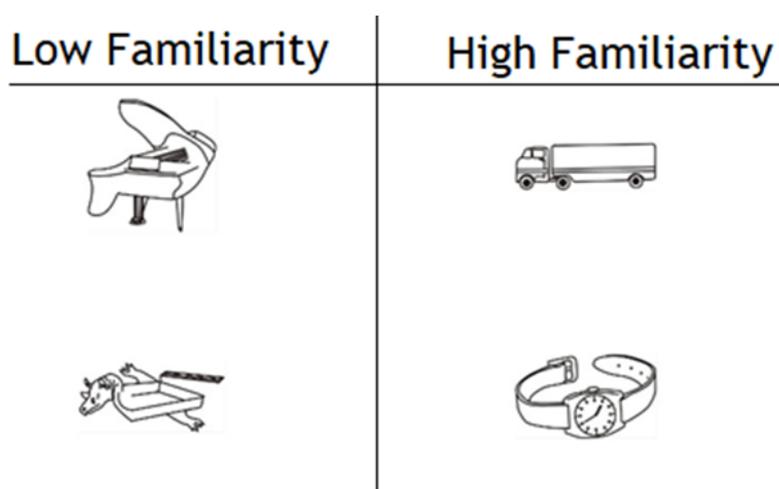


Figure 2: example of two low and two high familiarity images used in experiment 1

Procedure

This experiment used a complex span paradigm in which stimuli to remember were the familiar or the unfamiliar images described in the material section (figure 3). In addition to the memory task, a concurrent processing task was inserted after each image presentation. The concurrent task was a parity judgement task with digits ranging from 1 to 9, 5 being excluded. In addition to the image type, we also manipulated the delay between a mask directly following image presentation and the beginning of the secondary task pace.

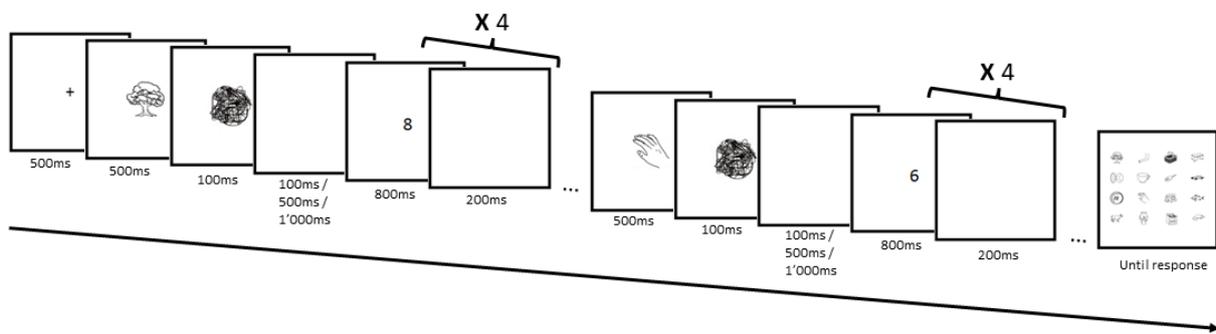


Figure 3: Schematic representation of a trial with two items presented in Experiment 1.

Every trial started with a screen informing the participant of the number of items to be remembered and reminding them to start performing articulatory suppression (repeating "BA-BI-BU" out loud). To start the trial, they had to press the space bar. After pressing it, a fixation cross appeared for 500ms in the middle of the screen. After the 500ms, the first image to memorize appeared on screen for 500ms. Just after the disappearance of the image, a mask stayed on screen for 100ms. The mask was created by the superposition of fourteen low familiarity images taken from the Soldan et al.'s image set but that were not used in this experiment. After the mask disappeared from the screen, a variable delay was inserted before the beginning of the processing task. The delay could last 100ms, 500ms or 1'000ms.

The processing task started after the delay. The processing task was a parity judgment task. Four digits were presented sequentially, each stayed on screen for 800ms and there were a 200ms blanks after each digit. Participants were tasked to judge whether the digit was odd or even, by respectively pressing the left or right

mouse button. Participants had to execute this judgement while the digit was on the screen or during the 200ms blank following the digit presentation. After the fourth digit disappeared from the screen and the associated 200ms blank, the next image to memorize appeared on screen. The image presentation and the associated processing phase would repeat for how many images there were to maintain in the trial.

After the last image presentation and processing phase, 16 images appeared on screen in a 4 x 4 matrix. The 16 images comprised of the presented items and other non-target images of the same type but not used in the current trial, randomly ordered. Participants were then tasked to click on the presented images in the order they were presented in. They could not correct themselves after clicking on an image, but they had no maximum time to answer. After the participant had clicked on his last response, the next trial started.

The list length of items to memorize varied from 2 to 5. The experiment was divided in three blocks, one for each consolidation delay. Inside each block, there were three trials per list length and image type for a total of 24 trials per block. The trials order within each block was randomized for each participant. The block order was counterbalanced across participants. The experiment had 72 trials in total.

The presented images were randomly selected from a pool of 42 images for each image type. The same pool of images was used in each of the three blocks. Inside each block, each image was used only once as a memorandum. The other non-target images in the response phase were drawn from another pool of items from the same type but that weren't used as memoranda at all. These images were used approximately 7 times each as non-target in each block.

In addition to the experimental trials, participants underwent a training specifically for the parity task and then three complete training trials with the "medium" consolidation delay and both kinds of images. In addition to these general training trials, the first three trials of each block were used as a training for the

consolidation delay of the block. These 3 training trials at the beginning of each block were not analyzed. The three general trainings and the three training trials at the beginning of each block used the same pool of item, which was not the same pool of items used in the experimental trials.

5.5.2 Results

First, we assessed participant’s performance to the concurrent task. All participants reached our inclusion criterion to have at least 70% of correct response to the parity judgement task. Then, we evaluated the impact of the familiarity manipulation and the consolidation manipulation on recall performance. To this aim, we computed the partial credit unit (PCU) for each experimental cell and each participant (Figure 4).

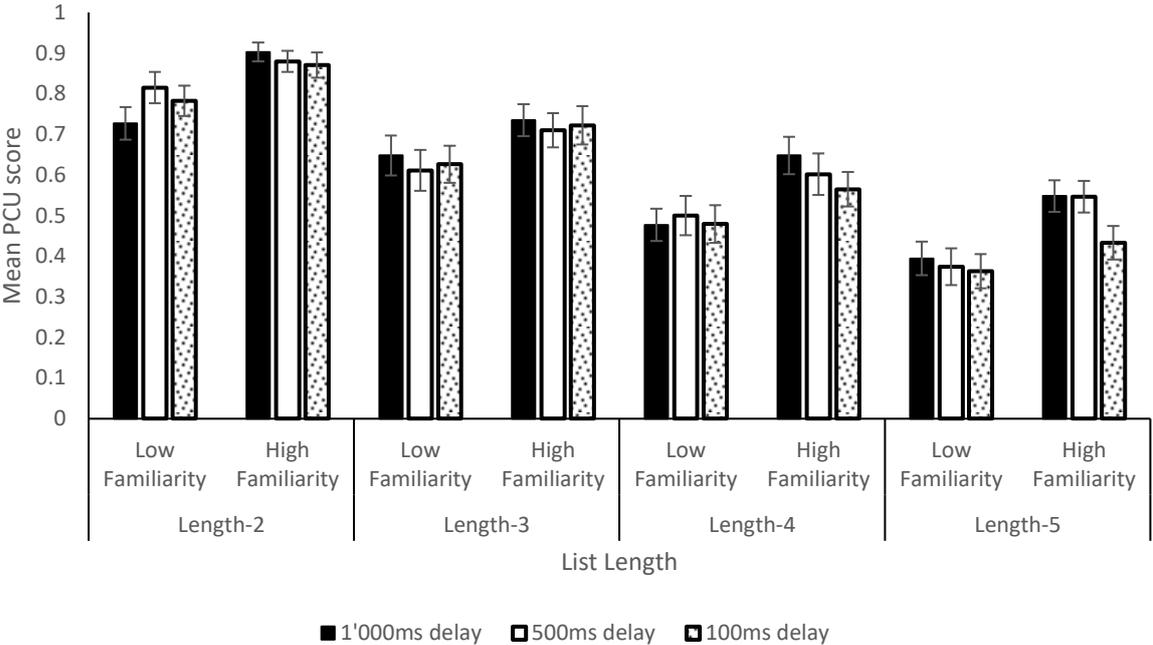


Figure 4: Mean PCU score in Experiment 1 as a function of image familiarity (Low or High), consolidation delay (100ms, 500ms or 1000ms) and list-length (2 to 5). Error bars are standard error.

The PCU is a score going from 0 to 1 and is computed as the mean proportion of images correctly recalled in the correct position in each trial. It has been shown to be a good index of recall performance in working memory span tasks (Conway, Kane, Bunting, Hambrick, Wilhelm & Engle, 2005). Since recall performance is influenced by

the total number of items presented (i.e., the list length), and we orthogonally manipulated list-length to the other variable, we included list-length as a predictive factor in the following analysis. We then used the PCU as dependent variable in a 2 (familiarity: high or low) \times 3 (consolidation delay: 100ms, 500ms or 1'000ms) \times 4 (list length: 2 to 5) repeated measure Bayesian ANOVA. The best model coming out of this Bayesian ANOVA was the model containing the main effect of image familiarity and list-length, $BF_{10} = 9.1 \times 10^{95}$, with high familiarity images yielding better recall performance than low familiarity images, and higher list-length inducing worse recall performance than lower list-length. There was overwhelming evidence for the inclusion of image familiarity ($BF_{inclusion} = 1.0 \times 10^{16}$, mean PCU score = 0.68 ± 0.18 and 0.57 ± 0.22 for high and low familiarity respectively) and for the inclusion of list-length ($BF_{inclusion} = 1.4 \times 10^{86}$, Mean PCU score = 0.83 ± 0.21 ; 0.68 ± 0.28 ; 0.55 ± 0.28 , and 0.44 ± 0.26 for list-length 2 to 5, respectively) There was evidence against an effect of consolidation delay, $BF_{exclusion} = 21$ (mean PCU score = 0.61 ± 0.21 , 0.63 ± 0.21 , and 0.63 ± 0.21 for 100, 500 and 1'000ms), and against the interaction of interest between consolidation delay and image familiarity, $BF_{exclusion} = 11$. Thus, recall performance was only influenced by image familiarity and list length, with evidence against an effect of consolidation delay and the interactions.

To assess whether more familiar memoranda were consolidated more quickly than less familiar memoranda, we assessed the effect of image familiarity and consolidation delay on the reaction times to the first parity judgment of each parity episode (i.e., Initial-RTs). The rationale was that because consolidation uses central attention, it will postpone any other attentionally demanding process until consolidation is completed. To ensure that images were consolidated, we only kept Initial-RTs in trials that were successful (all images recalled in correct position, 57.8% of trials excluded) and then computed the mean reaction time of each experimental cell and each participant. Since Initial-RTs have been shown to be influenced by the memory load (Camos et al. 2019, Vergauwe et al. 2014), we computed Initial-RTs as a function of memory load, image's familiarity and consolidation delay. It is important

to note that the memory load is not the same as the list length. The memory load consists in the number of images presented before the parity phase the reaction time is taken from. For example, in a trial with four memoranda (list-length of 4), there would be four distinct measures of Initial-RTs: one for each parity phase following each memoranda presentation. In contrast, in a trial with only two images presented, there would be two distinct measures taken. We then used the mean Initial-RTs as dependent variable in a 2 (familiarity: high or low) x 3 (consolidation delay: 100ms, 500ms or 1'000ms) X 5 (memory load: 1 to 5) repeated measure Bayesian ANOVA. The best model coming out this analysis showed an effect of consolidation delay and image type on Initial-RTs ($BF_{10}=51$), with convincing evidence for slower Initial-RTs in low familiarity images condition compared to high familiarity images condition, and slower reaction time induced by longer consolidation delay. Regarding the effect of memory load, we found evidence against its inclusion in the best model ($BF_{\text{exclusion}} = 5.8$). However, although we found ample evidence for the presence of the effect of image type ($BF_{\text{inclusion}} = 39$), there was only tentative evidence for the inclusion of consolidation delay ($BF_{\text{inclusion}} = 1.3$). Since there was evidence against the inclusion of the memory load factor in the model, we collapsed Initial-RTs as a function of image familiarity and consolidation delay and then used it as dependent variable in a 2 (familiarity: low or high) X 3 (consolidation delay: 100ms, 500ms or 1'000ms) repeated measure Bayesian ANOVA. This Bayesian ANOVA showed that the best model to account for the data included both effects of image familiarity and consolidation delay, $BF_{10} = 107$ compared to the null hypothesis (Figure 5).

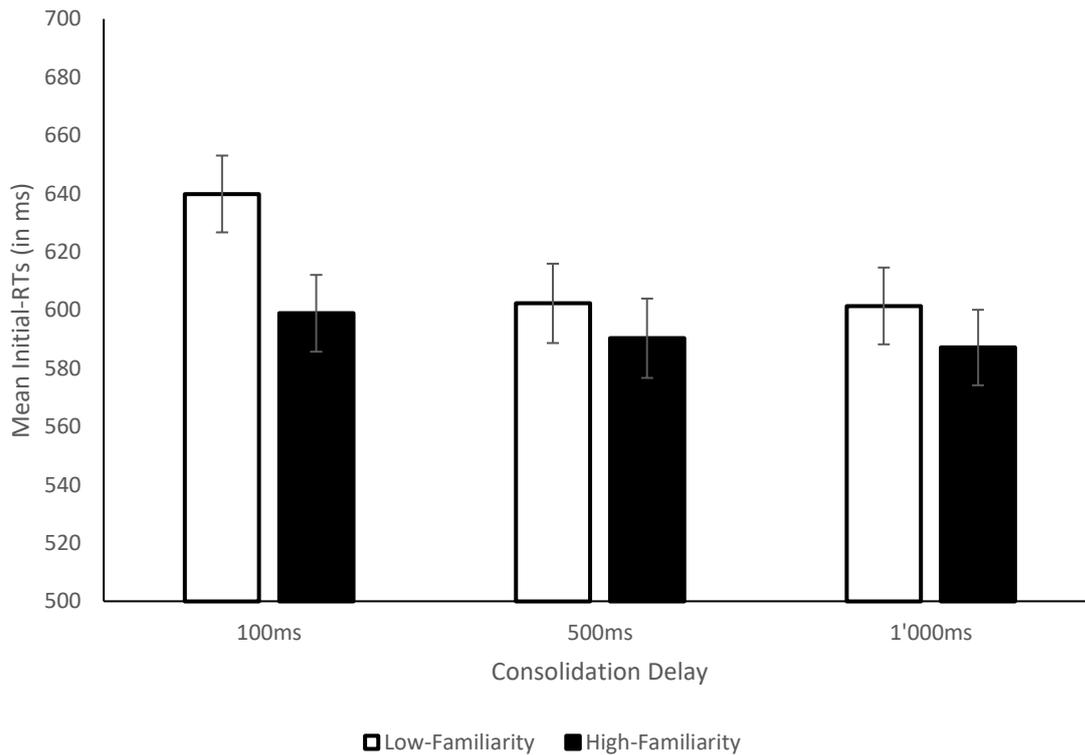


Figure 5: Mean Initial-RTs of every parity task phase, taken only from successful trial (all images recalled in the correct location), as a function of consolidation delay and image's familiarity. Error bars are standard error.

More familiar images yielded shorter reaction times, $BF_{inclusion} = 33.4$ (mean reaction time: $592\text{ms} \pm 79\text{ms}$ and $614 \pm 81\text{ms}$ for high and low familiarity images, respectively) and shorter consolidation delays yielded longer reaction time than longer consolidation delay, $BF_{inclusion} = 3.7$ (mean reaction time: $619\text{ms} \pm 81\text{ms}$; $596\text{ms} \pm 81\text{ms}$ and $594\text{ms} \pm 78\text{ms}$ for the 100ms, 500ms and 1'000ms consolidation delay, respectively). There was ambiguous evidence against the interaction, $BF_{exclusion} = 1.4$. Although the analysis on Initial-RTs showed evidence for an effect of consolidation delay and image familiarity on Initial-RTs in the predicted direction, we found weak evidence against the inclusion of the interaction.

Finally, we assessed whether the mean reaction time to the processing task (1st digit excluded, i.e., Subsequent-RTs) was impacted by our familiarity and consolidation delay manipulation. As with Initial-RTs, several studies showed that

Subsequent-RTs are also influenced by the memory load at the moment the processing episode takes place (Camos et al. 2019; Vergauwe, Camos, Barrouillet, 2014; Schneider, Vergauwe & Camos, in prep). Thus, we also included the memory load as a predictive factor in this analysis. We only kept trials where participants successfully recalled all images in the correct position. We computed the mean Subsequent-RTs as a function of memory load, consolidation delay and image familiarity and used it as dependent variable in a 2 (image familiarity: low or high) x 3 (consolidation delay: 100ms, 500ms or 1'000ms) x 5 (memory load: 1 to 5) repeated measure Bayesian ANOVA. The best model coming out of this analysis showed only a simple effect of memory load ($BF_{10} = 286$) with longer Subsequent-RTs induced by higher memory load (mean reaction time: 1 image = 502ms \pm 65ms; 2 images = 517ms \pm 65ms; 3 images = 518ms \pm 75ms; 4 images = 514 \pm 77; 5 images = 526ms \pm 90ms). The analysis also yielded evidence against an effect of consolidation delay, $BF_{\text{exclusion}} = 45$ (100ms delay = 511ms \pm 70ms; 500ms delay = 515ms \pm 77ms; 1'000ms delay = 513ms \pm 71ms) and against an effect of image familiarity, $BF_{\text{exclusion}} = 11$ (high familiarity = 514ms \pm 73ms; low familiarity = 513ms \pm 71ms), on Subsequent-RTs.

5.5.3 Discussion

The results from Experiment 1 are congruent with the idea that more familiar memoranda are consolidated more quickly than less familiar memoranda. This was exemplified by the pattern of results, showing that more familiar memoranda yielded shorter reaction time to the processing task than less familiar memoranda. However, although we found an effect of consolidation delay on Initial-RTs, there were evidence against an effect of this variable on recognition performance. It could be that the effect on Initial-RTs reflects the attentional cost induced by switching from the memory task to the concurrent task.

Switching between two tasks is costly and often apply a postponement in reaction time too (Kleinsorge & Gajewski, 2004; Wylie & Allport, 2000). Thus, shorter

consolidation delay would induce less time for task preparation, and this would elongate Initial-RTs. However, if the effect of consolidation delay on Initial-RTs were only due to task preparation, we would not expect an effect of image familiarity on Initial-RTs, for which we gathered evidence. Indeed, although task preparation has been shown to be influenced by characteristics of the preceding task (Los & Van Der Burg, 2010), only consolidation can explain the impact of memoranda familiarity on Initial-RTs. There is no reason that task preparation should be influenced by the familiarity of the memoranda. Thus, the impact of memoranda familiarity on Initial-RTs can be interpreted as more familiar memoranda being consolidated more quickly than less familiar memoranda.

The absence of consolidation delay impact on recall performance could be explained by the recognition procedure we used. Several other experiments investigating consolidation found an effect of SOA manipulation on reaction time to the concurrent task but not on recall performance (Jolicoeur & Dell'Acqua, 1997; Stevanovski & Jolicoeur, 2007). The impact of SOA manipulation on recall performance was only found when the concurrent task and the memory task were more difficult (Nieuwenstein & Wyble, 2014, Ricker & Cowan, 2014). Notably, studies that found an effect of SOA manipulation mostly used recall procedure (De Schrijver & Barrouillet, 2017; Nieuwenstein & Wyble, 2014), whereas studies that used recognition procedure were less likely to show an impact of SOA manipulation on recognition performance (Jolicoeur & Dell'Acqua, 1997; Stevanovski & Jolicoeur, 2007). This could be explained by the fact that recognition procedure can rely on a global matching process (see Clark & Gronlund, 1996 for a review). This supposes that successful recognition can rely on more degraded working memory representations, because response can be based on a "feeling" of familiarity, which will be more pronounced on presented memoranda, compared to non-targets.

This is in opposition to recall, where the working memory representation must withstand temporal decay and interferences to a higher degree to be effectively

recalled. Comparing impact of our manipulation on recognition performance and Initial-RTs to the concurrent task, our pattern of results could mean that, although more familiar items are consolidated more quickly than less familiar items, both kinds of memoranda are consolidated sufficiently enough to attain the same level of recognition performance. This would mean that there is a minimum level of consolidation that items must attend to be successfully recognized, and this level is lower for recognition procedure than recall procedure.

Experiment 2 was designed to address this last possibility by asking participants to recall memoranda instead of recognizing them. In addition, we manipulated familiarity more directly by training half of the participants on the figures used as memoranda, whereas the other half would discover them during the experimental session.

5.6 Experiment 2

This second experiment used a similar complex span task than Experiment 1, with a few differences. Firstly, the participants had to recall the presented memoranda in correct order instead of recognizing them in correct order. Secondly, we manipulated familiarity more directly than in Experiment 1 and it was manipulated between-subject. Half of the participants had to attend five training sessions during which they learned the set of memoranda used in the experiment before doing the experimental procedure. The other half had the exact same experimental procedure as the trained group, but without the training sessions. The idea was that participants that had the training session would have better long-term memory representation of the specific memoranda used in the experiment compared to the group of participants that did not attend the training sessions. This way, we could directly manipulate the difference in long-term memory representation between both groups of participants. The memoranda comprised of a set of 20 simple abstract figures (Figure 6).

Lastly, the memory load only varied from 1 to 4 memoranda to increase the number of trials that could be done during the experimental procedure. The experimental procedure comprised of the complex-span and of two different control tasks created to ensure that the training session did induce better long-term memory representations. To this aim, each of the control tasks was run twice, once before and once after the complex span. This was done to insure that, even after the complex span task, participants in the no-training group had still worse long-term memory representation than participants in the trained group, despite possible learning effect induced by the task itself.

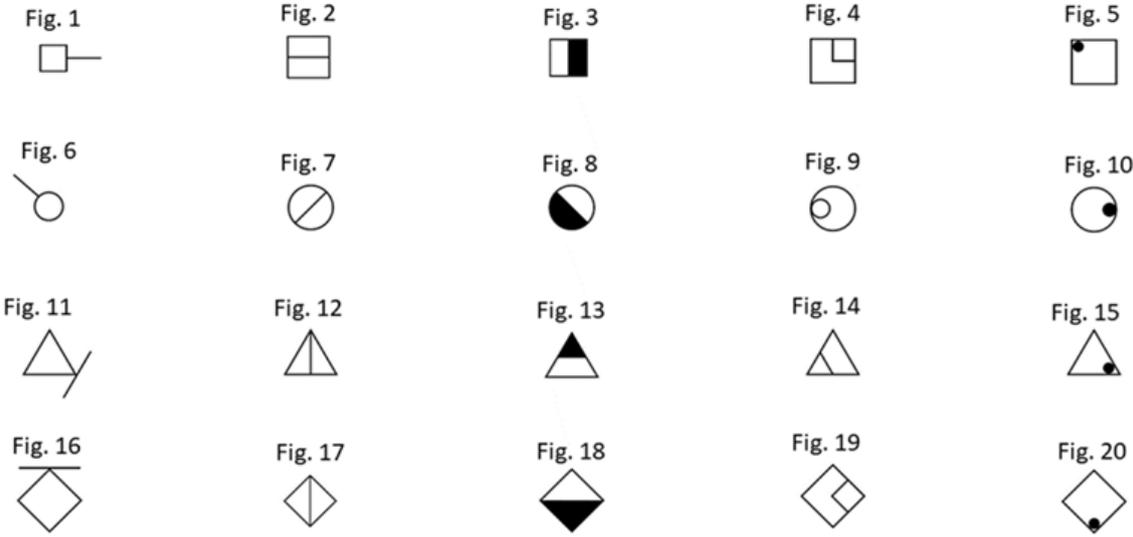


Figure 6: Complete set of abstract figures used in experiment 2. Each figure consists of the conjunction of a shape (square, circle, triangle, or diamond) and a feature (a connected segment, a dividing line, a colored portion, an included shape, or a dot).

The first control task was a timed copy task where each figure was presented on screen sequentially and participants had to redraw all the presented figures as quickly as possible. The second control task was a timed free recall where participants had two minutes to recall as many figures as possible that were presented in the timed copy task. The idea being that the trained group should be at ceiling in free recall placed before and after the complex-span (hereafter: pre- and post- complex span task, respectively), whereas the no-training group should improve the recall score in between the pre- and post- complex-span. Regarding the timed copy task,

trained participants should be quicker than untrained participants, and both groups should be quicker in the post complex-span task compared to the pre-complex span task.

5.6.1 Methodology

Participants

Seventy-seven students from the University of Fribourg (68 females, mean age: 20.9 ± 2.4 years) participated in this experiment. They were remunerated with partial course credits or a 15CHF (16.40USD) voucher. Every participant read and signed a form of consent. Ethical approval was given by IRB of the University of Fribourg.

Storage material

The memory items consisted of a set of 20 abstract figures (Figure 6). Each figure was created by the association of a basic shape (square, circle, triangle, or diamond), a feature (a connected segment, a dividing line, a colored portion, an included shape, or a dot) and the location of that attribute in the shape (Barrouillet, in prep).

Procedure

This experiment used a complex span paradigm in which stimuli to remember were the abstract figures described above. In addition to the memory task, a parity judgement task (on digits ranging from 1 to 9, 5 being excluded) was inserted after each image presentation. We manipulated participants familiarity to the abstract figures (between-subject manipulation) as well as the delay between item presentation and the beginning of the parity judgment task (within-subject manipulation, hereafter: consolidation delay). The number of memory items varied from 1 to 4.

Training sessions

To manipulate the familiarity of the figures used as memory items, half of the participants were randomly assigned to a training group in which they were familiarized with the abstract figures. To this end, participants in the training group had to attend five training sessions, once a day for one week, from Monday to Friday. Training sessions happened online and lasted approximately 15 minutes each. During them, each of the 20 abstract figures was presented one by one by the experimenters using the "share screen" function of the videoconference program we used.

Participants had to redraw each figure five times on a blank paper sheet. Each figure stayed on screen until all participants finished to redraw it. The figures were always presented in the same order. During the first training session, participants only had to redraw the presented figures. Starting from the second training session, participants had two minutes to recall a maximum of abstract figures before the training started. They recalled the figures on the back of the blank sheets they would use to do the training task. This free-recall allowed to assess how learning evolved across sessions. Participants had to bring the sheets on which they did the training sessions when they came for the experimental session so we could analyze their training performance. Since we wanted to control the timing between training sessions and the experimental session, participants in the trained group did the experiment only on Monday and Tuesday following the week they had the training sessions. The other half of the participants (the untrained group) did only the experimental procedure described below.

The experimental session

The experimental session was divided in five parts. In the first part, called "timed copy", participants were presented with all the abstract figures and had to recopy them as quickly as possible. Each figure was presented one by one on the computer screen. Participants were tasked to recopy the figure on a piece of paper, using a pen, and then press the space bar to show the next one. There was a 500ms

blank screen following the space bar press. We measured the time in between the appearance of each figure on screen and when participant pressed the space bar to show the next figure. This duration was used as a proxy of the effectiveness of the training manipulation. In addition, this phase would allow participants in the no-training group to see and copy each figure at least once before the complex span task. After they finished to copy the figures, the second part of the experiment started. It was a free recall of the figures that were just presented. Participants had two minutes to recall as many figures that were just presented as possible.

The third part of the experiment involved the complex-span task. Every complex-span trial started with a screen informing the participants of the number of items to memorize and consolidation delay for the upcoming trial, to put their dominant hand on the mouse and to press the space bar to start figures presentation. After they had pressed the space bar, a fixation cross appeared in the middle of the screen for 500ms and was followed by the first figure to memorize, which stayed on screen for 500ms. Right after it disappeared from the screen, the item was replaced by a mask, which stayed on screen for 100ms. The mask consisted of a superposition of 14 of the abstract figures which were not used as memoranda in the complex-span task. The mask was followed by a blank screen of either 100ms (short consolidation delay) or 1'000ms (long consolidation delay).

After this delay, the parity judgment task started. The parity task consisted of 4 digits sequentially presented. Each digit stayed on screen for 800ms, followed by a blank screen during 200ms before the next digit would appear on screen. Participants were tasked to judge whether the digit was odd or even while the digit was on-screen. To this aim, they had to press the left or right mouse button if it was respectively odd or even. After the 200ms blank following the fourth digit, the next item to memorize would appear on screen. The alternance between item presentation and parity judgment task would repeat for as many memory items there were in the trial.

After the last digit following the last item presented for the trial, a question mark appeared in the middle of the screen. Participants then had to recall the presented memory items in the order they were presented by drawing them on a piece of paper provided by the experimenter. The response paper comprised of one line for each trials and four columns. The first figure had to be recalled in the first column, the second figure in the second column, and so on. They were tasked to put a question mark in columns where they did not remember the answer. After they recalled every figure of the trial, participants had to press the space bar to finish the trial and make the information about the next trial appear on screen.

The number of items to maintain varied from 1 to 4 figures. There were 6 trials per list length and consolidation delay, for a total of 48 experimental trials. The trial order was pseudo-randomized: two trial lists were created, and each participant were randomly assigned one of the lists. Each list was constructed as follows: each trial could contain only one exemplar of each shape (square, circle, triangle, or diamond). The list lengths were evenly distributed across the whole list (same number of 1-, 2-, 3- and 4- items trials in both halves of the list). There were no more than two consecutive trials in the same consolidation delay condition and there were the same number of long and short trials in both halves of the list. This was done so the different conditions were equally distributed along the experiment. Each abstract figure was used exactly 6 times as memoranda (3 time in the first half and 3 time in the second half of the experiment) and none were used in two consecutive trials. There were 6 trials per experimental cells, 3 in the first half of the experiment and 3 in the second half of the experiment.

Before the first block, participants had two trainings: one to the parity judgment task alone and then a general training to the complex-span task with 4 trials, two with the long consolidation delay and two with the short one. For both consolidation delay, there were one training with 1 item to maintain and the other with 2 items to maintain. For the training to the parity task, 20 digits appeared

sequentially on screen at the pace described before. Participants had to sort them while the digits stayed on screen.

After they finished the complex-span task, participants did the fourth and the fifth part of the experiment, which were the same as the first and second part respectively. For the fourth part, directly after the end of the complex span task, participants had two minutes to recall a maximum of figures that were used in the experiment (i.e., free recall post-complex span). Finally, for the fifth part, after this second free recall, they had to do a last timed copy of the figures, which followed the same procedure than the first timed copy.

5.6.2 Results

First, we assessed the ratio of correct parity judgement for each participant. Participants with less than 70% of correct responses to the parity task were discarded from the analysis. This led to the exclusion from the analysis of the data of 8 participants (4 in each group). The final pool had a total 69 participants (9 men), 41 (4 men) in the no-training group and 28 (5 men) in the trained group.

Analysis free-recall during training sessions

Before analyzing results to the experimental session, we analyzed the free recall score during the training sessions. The free recall was scored as follows: Each figure could give between 0 and 2 points. Two points were given if the participants could recall the correct shape containing the correct feature in the correct location, and one point if the correct shape contained an incorrect feature (wrong location of the feature or wrong feature). Finally, no point was given in all other situations. The recall score for each participant in each free recall consisted in the total number of points earned by the participants divided by 40 (maximum amount of point gainable). The recall score thus consisted of a number between 0 and 1, where 1 meant perfect recall of all figures and 0 meant no figures at all were recalled. To minimize the subjectivity of the rating, three experimenters rated the free recall independently from each

other. The score coming from each experimenter correlated between each other from 0.81 to 0.95, which show a high level of agreement between experimenters.

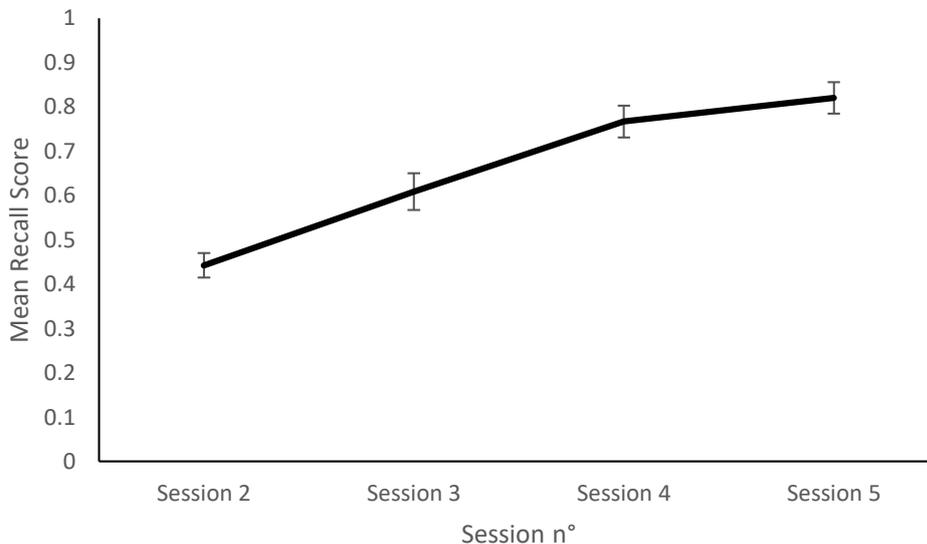


Figure 7: Mean recall score in the training's free-recall as a function of training session in experiment 2. Error bars are standard errors.

We then used this score as a dependent variable in a one-way repeated measure Bayesian ANOVA with training session (2 to 5) as within subject factor (Figure 7). This analysis yielded extreme evidence for an effect of training sessions, $BF_{10} = 4.6 \times 10^{17}$. Recall score increased with each new session, showing that participant did learn the figures.

Analysis experimental free recall

Before analyzing results from the complex-span task, we assessed whether the trained group effectively had better long-term memory representation than the no-training group. To this aim, we first analyzed recall scores to the experimental free recall as a function of task location (pre- or post- complex-span) and training group (trained or no-training). The experimental free-recall score was computed exactly the same way as the free-recall score from the training sessions. The score coming from each experimenter correlated between each other from 0.92 to 0.97. Once computed, we used this score as dependent variable in a 2 (training group: trained or untrained) X 2 (task location: pre- or post- complex-span) mixed Bayesian ANOVA. The best

model coming out of this analysis was the full model, with evidence for an effect of training group, task location and the interaction between both factors, $BF_{10} = 1.3 \times 10^{28}$ (Figure 8).

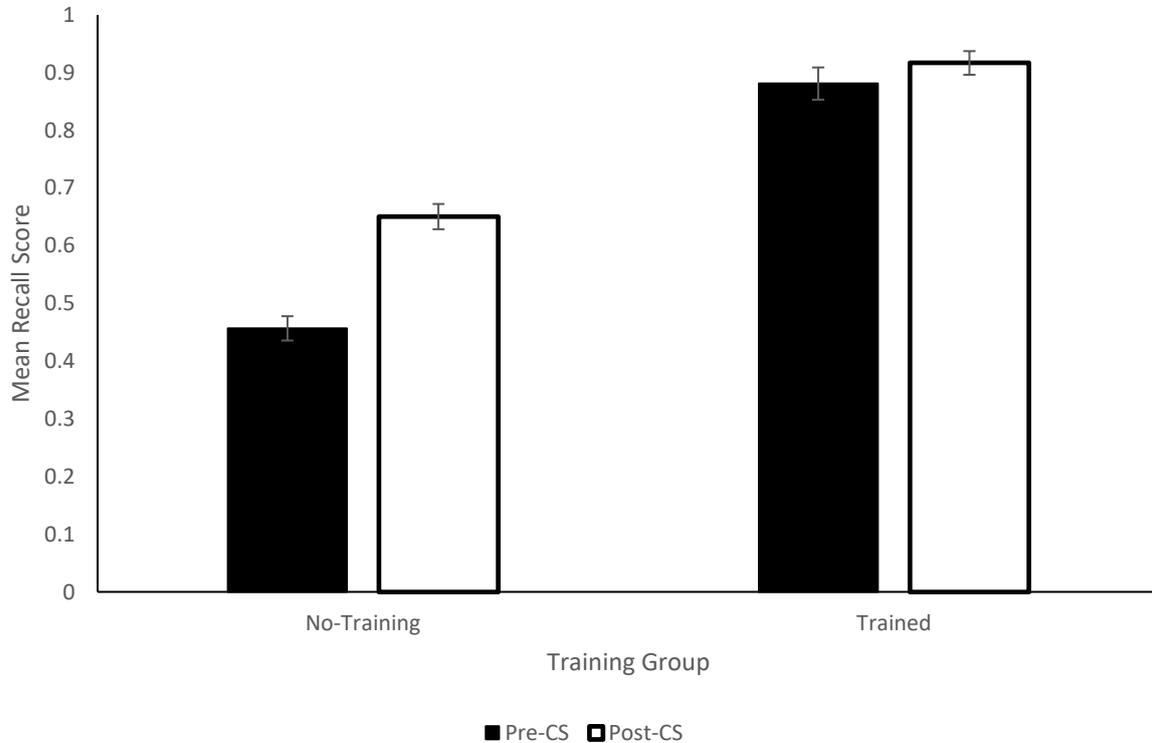


Figure 8: Recall score to the free recall as a function of training group (trained or no-training) and task location (pre- or post- complex-span). Error bars are standard error.

Participant in the trained group had better recall score than the untrained group, $BF_{inclusion} = 1.6 \times 10^{14}$ (mean score = 0.90 ± 0.13 and 0.55 ± 0.17 for the trained and untrained groups) and participants were better in the post- than in the pre-complex span location, $BF_{inclusion} = 6.4 \times 10^8$ (mean score = 0.63 ± 0.25 and 0.76 ± 0.18 for pre- and post-complex span). The interaction showed that difference between groups was larger before the complex span than after it, $BF_{inclusion} = 1.8 \times 10^5$. Results from the experimental free recall show that participants in the no-training group recalled less figures than participants in the training group. The no-training group had a bigger enhancement between both free recall than the trained group, but this is due to the trained group being almost at ceiling in both free recall. Notably, participants in the no-training group did not attend the same level of performance than the trained

group, showing that the complex-span did not induce as well a training on the figures compared to the training sessions.

Analysis timed copy

The second assessment of a difference in long-term memory representation between both training groups relied on the analysis of the time participants took to copy each image once. We summed the total time, in seconds, it took for each participant to do the timed copy and use this variable as a dependent variable in a 2 (training group: trained or no-training) x 2 (task location: pre- or post-complex-span) mixed Bayesian ANOVA (Figure 9). The best model coming out of this analysis was

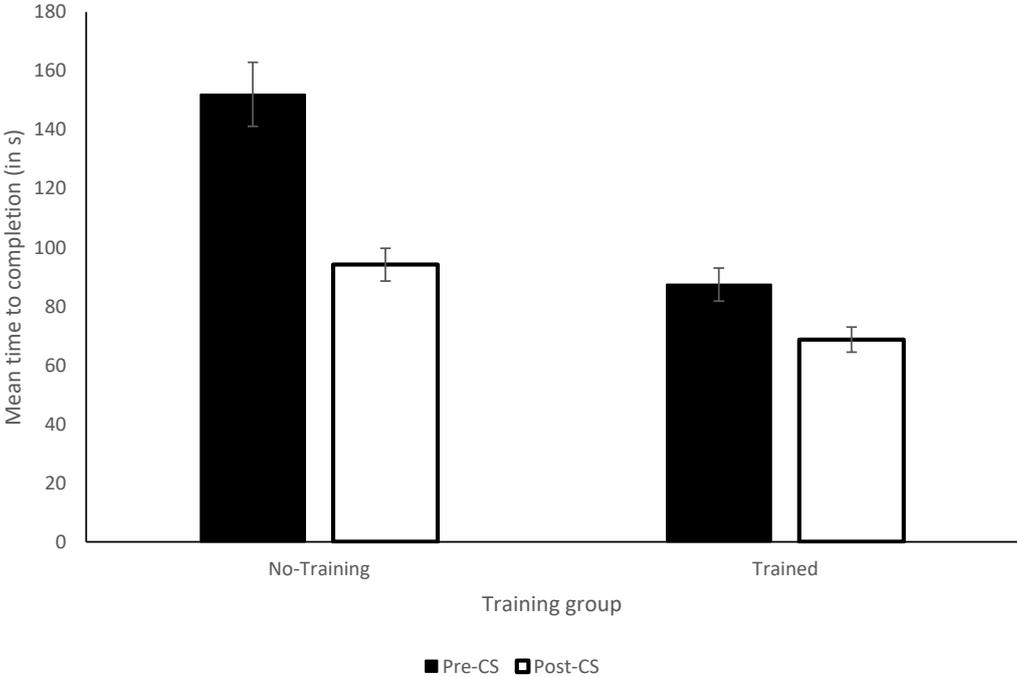


Figure 9: Mean time to completion to the timed copy task (in second) as a function of training group (trained or no-training) and task location (pre- or post- complex span). Error bars are standard error.

the full model, which included an effect of training group, of task location and of the interaction between both factors, $BF_{10} = 5.1 \times 10^{14}$. Participants in the trained group were faster than participants in the no-training group, $BF_{inclusion} = 351$ (Trained group = $78s \pm 28s$; no-training group = $123s \pm 62s$). Participants were also quicker in the post-complex span location compared to the pre-complex span location, $BF_{inclusion} = 1.9 \times 10^9$ (Time pre-complex span = $126s \pm 65s$; Time post-complex span = $84s \pm$

33s). The interaction indicated that the no-training group had bigger difference between the pre- and post- complex span location compared to the trained group, $BF_{inclusion} = 584$. The results show that participants in the no-training group were always slower than the trained group, and the no-training group never attained the same speed as the trained group. The results from the timed copy are congruent with the results from the experimental free recall, and both argues in favor of a better long-term memory representation for the trained group compared to the no-training group regarding the abstract figures.

Analysis of the complex span task

Now that it has been demonstrated that participants in the training group had better long-term memory representation of the trained shape than the no-training group, we turned to the analysis of recall score and reaction time to the concurrent task in the complex-span task as a function of training group and consolidation delay. First, we computed the recall score as follow: each image could give from 0 to 2 points. 2 points were given if the correct shape containing the correct feature in correct location was recalled in the correct serial position. One point was given if either the correct shape or the correct feature was recalled in the correct serial position. All other possibilities yielded 0 point. We then summed up the total score for each trial and divided it by the number of presented figures in the trial times two, because the total possible amount of point for each trial was the number of presented items times two. This way, a recall score between 0 (no points at all) and 1(all correct shapes and features recall in the correct location) was computed for each trial. To minimize possible subjective judgements, three experimenters rated each participant independently. Correlation between experimenter rating was between 0.95 and 0.98, showing a high degree of agreement between experimenters.

We then used the recall score as a dependent variable in a 4 (total number of images presented in the trial: 1 to 4) x 2 (Training group: trained or untrained) x 2 (consolidation delay: 100ms or 1'000ms) mixed Bayesian ANOVA (Figure 10). This

analysis showed that the best model to account for our data included two simple effects of the number of images and consolidation delay in addition to the interaction between both factors, $BF_{10} = 2.5 \times 10^{146}$. There was convincing evidence for an effect of consolidation delay, with shorter consolidation delay inducing worse recall performance than longer consolidation delay, $BF_{inclusion} = 236$ (Score short delay = 0.57 ± 0.28 ; Score long delay = 0.63 ± 0.23).

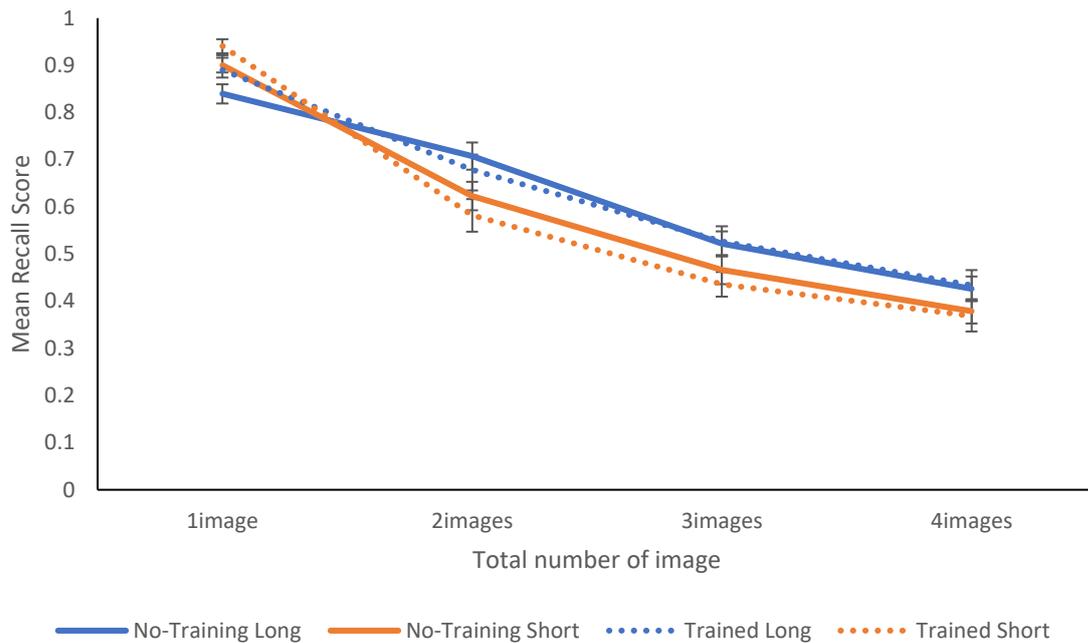


Figure 10: Mean recall score to the complex-span as a function of total number of image in the trial, training group and consolidation delay. Error bars are standard error.

The number of images presented in the trial also had an impact on recall performance, with worse recall performance the more figures there were to remember, $BF_{inclusion} = 1.6 \times 10^{139}$ (score 1 image = 0.89 ± 0.11 ; score 2 images = 0.65 ± 0.20 ; score 3 images = 0.49 ± 0.17 ; score 4 images = 0.40 ± 0.17). There was extreme evidence for the presence of the interaction, $BF_{inclusion} = 1.2 \times 10^5$. Interestingly, there was evidence against an effect of training condition, $BF_{exclusion} = 4.0$ (score trained group = 0.61 ± 0.25 ; score untrained group = 0.61 ± 0.24). Other interactions were not evaluated, because they included a main effect of training group, and we gathered evidence against its inclusion in the best model. This analysis showed that recall score was only impacted by consolidation delay and memory load,

and not by our training manipulation. Higher memory load induced worst recall performance compared to lower memory load, and recall scores were higher when there was a longer consolidation delay. The presence of the interaction showed that the effect of consolidation delay on recall performance was more pronounced in trials with higher memory load.

To analyze the impact of consolidation manipulation on the concurrent task reaction time, we extracted Initial-RTs of each parity phase and analyzed it separately from the other reaction time (i.e., Subsequent-RTs). We then used the mean Initial-RTs as dependent variable in a 4 (memory load: 1 to 4) x 2 (training group: trained or no-training) x 2 (consolidation delay: short 100ms or long 1'000ms) mixed repeated measure Bayesian ANOVA (Figure 11).

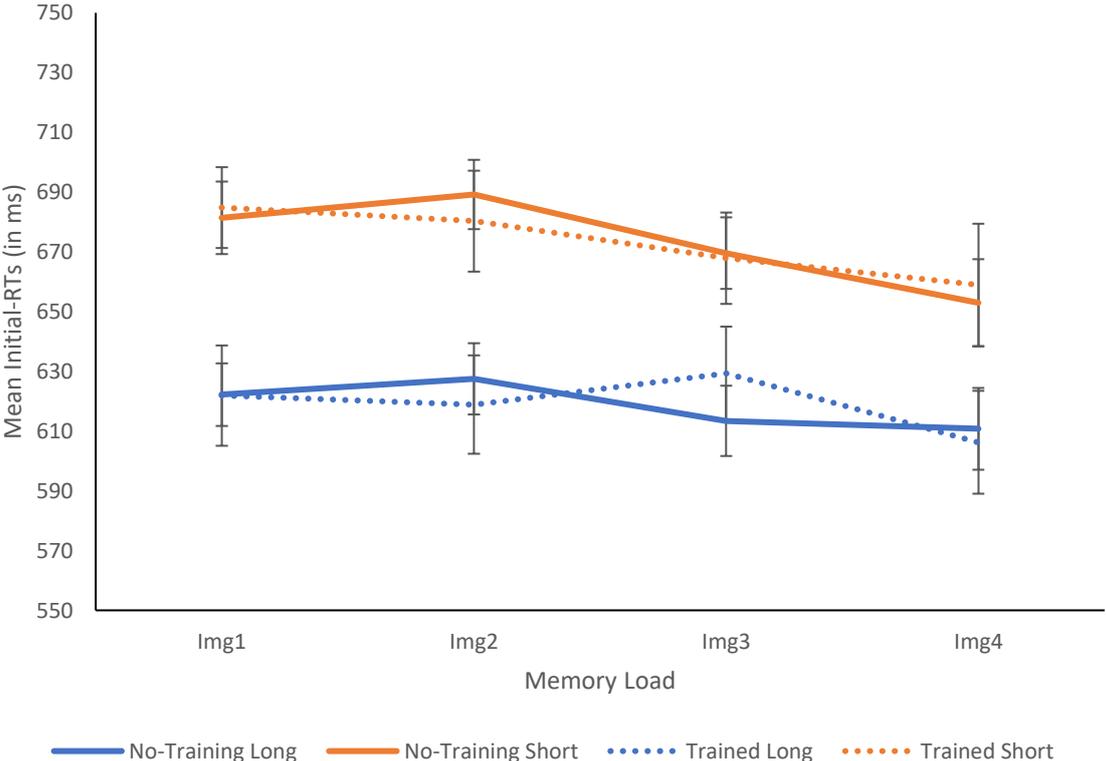


Figure 11: Mean Initial-RTs (in ms) of every parity phase as a function of memory load, training group and consolidation delay. Error bars are standard error.

We included memory load as a predictive factor for the same reason as in Experiment 1. The analysis showed that the best model to account for our data included two

simple effect of consolidation delay and memory load, $BF_{10} = 7.4 \times 10^{28}$. There was evidence for an effect of memory load, with higher memory load linked to quicker reaction time, $BF_{inclusion} = 14$ (1 image = 652ms \pm 83ms; 2 images = 655ms \pm 82ms; 3 images = 644ms \pm 88ms; 4 images = 632ms \pm 93ms) and evidence for an effect of consolidation delay, with slower reaction time induced by shorter consolidation delay, $BF_{inclusion} = 6.1 \times 10^{27}$ (short consolidation delay = 673ms \pm 87ms; long consolidation delay = 619ms \pm 78ms). We found evidence against an effect of training manipulation, $BF_{exclusion} = 3.1$ (Trained group = 646ms \pm 91; No-training group = 646ms \pm 83). We also gathered evidence against the inclusion of the interaction between memory load and consolidation delay manipulation, $BF_{exclusion} = 21$. We did not evaluate the other interactions, as they included the effect of training group, for which we gathered evidence against its inclusion. This analysis showed that Initial-RTs were influenced by the consolidation delay and the memory load, but not by our training manipulation. Initial-RTs were slower with shorter consolidation delay, but each presented item increased the speed at which participants answered to the Initial-RTs. Interestingly, we found evidence against an effect of training manipulation. Participants in the no-training had similar Initial-RTs than participants in the trained group in all conditions.

Finally, as in Experiment 1, we analyzed the impact of memory load, training group and consolidation delay on the Subsequent-RTs. We extracted this mean in the same way as the Initial-RTs and used it as dependent variable in a 4 (memory load: 1 to 4) \times 2 (training group: trained or no-training) \times 2 (consolidation delay: short or long) mixed Bayesian ANOVA (Figure 12). This analysis showed that the best model to account for our data included only a simple effect of memory load, $BF_{10} = 2.1 \times 10^4$. There was evidence for an increase of the reaction time with higher memory load, $BF_{inclusion} = 2.2 \times 10^4$ (1 image = 549ms \pm 52ms; 2 images = 566ms \pm 54ms; 3 images = 567ms \pm 57ms; 4 images = 561ms \pm 64ms), but we found tentative evidence against an effect of training manipulation, $BF_{exclusion} = 2.8$ (trained group = 563ms \pm 59ms; No-training group = 559ms \pm 56ms) and against an effect of consolidation

delay, $BF_{\text{exclusion}} = 4.3$ (Short consolidation delay = $559\text{ms} \pm 60\text{ms}$; Long consolidation delay = $563\text{ms} \pm 55\text{ms}$).

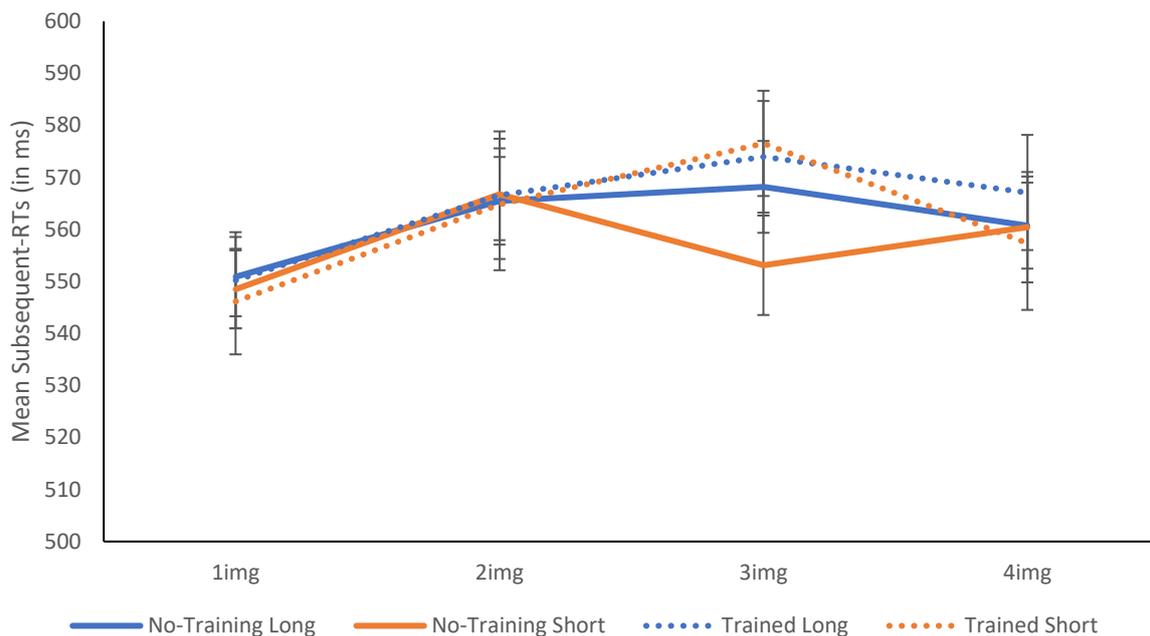


Figure 12: Mean Subsequent-RTs (in ms) as a function of memory load, training group and consolidation delay. Error bars are standard error.

This analysis showed that Subsequent-RTs were only influenced by the memory load, with longer Subsequent-RTs induced by higher memory load. There was tentative evidence for an absence of effect of training manipulation and substantial evidence against an effect of consolidation delay on Subsequent-RTs.

5.6.3 Discussion

Analysis of the pre- and post- complex-span free recall and timed copy showed that participants in the trained group had better long-term memory representations than participants in the no-training group. The no-training group had slower copy time than the training group and had worse recall score during both free recall tests, compared to the trained group. In addition, the interaction between task location (pre- and post- complex span) and training group (Trained or No-training) was present in both tasks. This is likely due to the fact that participants in the trained group were already at ceiling, or close to it, in the pre-complex span task, and thus

had limited possibilities to enhance their performance compared to the no-training group. Importantly, the no-training group never attained the same level of performance than the trained group, which solidifies the conclusion that the trained group effectively had better long-term memory representation of the figures than the no-training group.

In the complex span task, we found evidence of an effect of consolidation manipulation on recall performance in the predicted direction. Shorter consolidation delays yielded worse recall performance than longer consolidation delays. We also found the expected effect of consolidation manipulation on Initial-RTs. Longer consolidation delay induced quicker Initial-RTs compared to shorter consolidation delay. Interestingly, we found evidence against the effect of the training manipulation. There was no difference between the trained group and the no-training group on Initial-RTs and recall performance. This is a striking pattern, as the post- and pre- complex-span free recall task and timed copy showed that both groups differed on their familiarity with the material; the trained group having better long-term memory representation than the no-training group. Indeed, several other experiments showed that image familiarity influence recall performance, where more familiar images are better recalled than less familiar ones (Curby, Glazec & Gauthier, 2009; Jackson & Raymond, 2008). Possible explanations for this discrepancy are presented in the general discussion below (chapter 5.7.2).

Analysis of the Initial-RTs in the complex-span showed that this variable was influenced by the memory load as well as the consolidation delay manipulation. Initial-RTs were slower in trials where the consolidation delay was shorter. This was the expected pattern of results, as this effect was at the basis the consolidation literature. Taken alone, the effect of consolidation delay on Initial-RTs could also be explained by task preparation, as in Experiment 1. However, the fact that we also found an effect of consolidation delay on recall performance argues against this possibility. There is no theoretical reason to think that task preparation, in itself,

should have an impact on recall performance. The impact of consolidation delay on recall performance is better explained by the involvement of a consolidation process. In trials with shorter consolidation delay, Initial-RTs were postponed until consolidation process was over, whereas, in trials with longer consolidation delay, consolidation process was already over when the first digit appeared. Initial-RTs were thus not postponed until consolidation was over.

Regarding the effect of memory load on Initial-RTs, although an effect was expected, previous experiments found that the effect pointed in the opposite direction: Higher memory load induced slower Initial-RTs (Camos et al. 2019; Vergauwe et al. 2014). In our results, higher memory load induced quicker Initial-RTs. One possible explanation for this discrepancy could rely on the fact that Camos and colleagues (2019) and Vergauwe and colleagues (2014) used a Brown-Peterson task to investigate the impact of memory load on RTs to a concurrent task, whereas we used a complex-span task in our experiments.

In a Brown-Peterson task, all memoranda are presented one after the other and a concurrent task appears after all memoranda have been presented, before the recall phase. Thus, only one memory load can be investigated during each trial. In contrast, a complex-span task allows to sample several Initial-RTs at different memory load during one trial. Although it allows to have more data points for the same overall experiment duration, the drawback is that higher memory loads are always at the end of a trial, whereas lower memory loads are always at the beginning of a trial. It may be that participants were more efficient to switch from encoding the memoranda to processing the concurrent task at the end of a trial compared to the beginning of a trial, because of task specific training effects. If this is the case, then participants had shorter Initial-RTs with higher memory load, because they were more trained to the concurrent task at this point compared to the beginning of a trial.

Finally, regarding Subsequent-RTs, we found only an effect of memory load on this variable. Higher memory load induced longer mean Subsequent-RTs. This last

result was to be expected, as this effect has been observed in several complex-span experiments and has been used as evidence for the existence of attentional refreshing. Attentional refreshing is a domain-general maintenance process that uses central attention to protect working memory representation against temporal decay and interferences (see Camos et al. 2018 for a review). All representations currently held in working memory are thought to be refreshed between processing episodes that also use central attention. Since attentional refreshing occupies central attention, no other attentionally demanding task can happen before refreshing is over. Thus, the more items are currently held in working memory, the longer refreshing of all of them will take, postponing the mean reaction time to the concurrent task (Vergauwe, Camos & Barrouillet, 2014; Camos et al. 2019).

5.7 General Discussion

In a series of experiments, we investigated whether consolidation was influenced by semantic long-term memory. More precisely, our hypothesis was that more familiar memoranda would be consolidated more quickly than less familiar memoranda. To this aim, we ran two experiments that relied on a complex-span paradigm.

In Experiment 1, we used a complex-span task in which we manipulated a consolidation delay orthogonally to memoranda familiarity, and we measured recognition score as well as reaction time to the concurrent task. Consolidation delay was operationalized as the delay between a mask directly following image presentation and the beginning of the concurrent task. We found evidence for an effect of consolidation delay and memoranda familiarity on Initial-RTs (first RT to each concurrent task phase). More familiar memoranda induced quicker Initial-RTs, and shorter consolidation delay induced slower Initial-RTs. However, contrary to our hypothesis, we found evidence against an effect of consolidation delay on recognition performance.

In Experiment 2, we also used a complex-span, but this time we manipulated memoranda familiarity in a between-subject experimental plan. Half of participants (the trained group) were trained on the specific figures used as memoranda before the experiment, whereas the other half of the participants (the no-training group) discovered the figures during the experiment. We controlled that participants in the trained group had effectively better long-term memory representation of the figures by including a timed-copy and a free recall task before and after the complex-span.

These pre-and post- complex-span tasks showed that participants in the trained group had effectively better long-term representation of the presented figures. Consolidation delay was operationalized in the same way as in Experiment 1. In addition, participants were tasked to redraw the presented figures, instead of simply recognizing them in correct order like in Experiment 1. Results showed that shorter consolidation delays induced slower Initial-RTs and worse recall performance than longer consolidation delays. However, we found evidence against an effect of our training manipulation. Initial-RTs, Subsequent-RTs and recall performance in the complex-span were statistically identical between the trained and the no-training group.

5.7.1 The effect of post-mask delay manipulation

In both experiments, we found the expected effect of post-mask delay manipulation on Initial-RTs. Shorter consolidation delay induced slower Initial-RTs. However, the effect of our consolidation manipulation on memory performance was present only in Experiment 2. In Experiment 1, where we found no evidence for an effect of consolidation delay on recognition performance, the effect could be explained as well by a consolidation process or by the attentional cost induced by the switching from the memory task to the concurrent task (Kleinsorge & Gajewski, 2004; Wylie & Allport, 2000).

In trials with longer consolidation delay, participants had more time to prepare for the concurrent task, compared to trials with shorter consolidation delay.

Participants could thus have responded more quickly to the concurrent task in trials with longer consolidation delay, because of task preparation. However, although contrary to our hypothesis, the absence of an effect of consolidation delay on recognition performance in Experiment 1 does not impede the idea that consolidation was effectively manipulated in this experiment. The first experiments investigating consolidation in working memory showed an effect on reaction time to a concurrent task without finding an effect on recognition performance (Jolicoeur & Dell'Acqua, 1997). The impact of consolidation delay on recognition performance was only evident when more difficult memoranda and concurrent task were used (Ricker & Cowan, 2014, Nieuwenstein & Wyble, 2014). It may be that our memory task was not difficult enough for an effect of consolidation delay to be apparent on recognition performance.

In line with this last proposition, Experiment 2 showed that Initial-RTs as well as recall performance were impacted by our consolidation delay manipulation. Shorter consolidation delay induced slower Initial-RTs and worse recall performance than longer consolidation delay. Although the effect of consolidation delay on Initial-RTs in Experiment 2, taken alone, could also be explained by the attentional switch cost, the fact that we also had an impact on recall performance argues in favor of an effect of consolidation instead. The effect on the memory task performance shows that our consolidation manipulation also impacted the maintenance of information in working memory. This effect on recall performance is straightforwardly explained by a consolidation process but is less well explained by the attentional switch cost hypothesis. Thus, results from Experiment 2 show that our post-mask delay manipulation effectively impacted consolidation in working memory, which is in line with former studies that investigated consolidation via a similar delay manipulation (Nieuwenstein & Wyble, 2014; De Schrijver & Barrouillet, 2017).

5.7.2 Does memoranda familiarity impact consolidation speed?

Results from Experiment 1 are congruent with the hypothesis that more familiar memoranda are consolidated more quickly than less familiar memoranda. This is evidenced by the impact of memoranda familiarity on Initial-RTs in our experiments. Participants responded more quickly to the first digit of every concurrent task phase in trials with more familiar memoranda compared to less familiar ones. If we consider that consolidation induces a processing bottleneck, where processing of other attentionally demanding task cannot be done simultaneously to the consolidation process, then the effect of memoranda familiarity on Initial-RTs can be explained. Indeed, the supposition is that more familiar memoranda are consolidated more quickly than less familiar ones. Consolidation process is over more quickly for familiar memoranda, and thus participants can switch to processing the concurrent task faster than in trials with less familiar memoranda.

In contrast, we found evidence against an impact of our familiarity manipulation in the complex-span of Experiment 2. Participants that were more familiar with the memoranda did not have statistically better recall performance or shorter Initial-RTs than participants that were unfamiliar with the memoranda. This is a striking result, as participants in the trained group had better long-term memory representation than the no-training group, as indexed by the pre- and post-complex-span tasks. Indeed, memoranda familiarity has been shown in the past to impact memory performance (Engle, Nations & Cantor, 1990; Morton, 1979; Clutterbuck & Johnston, 2002). There are several theoretical possibilities to explain our pattern of results and its discrepancy with the literature on familiarity.

One possible explanation for the discrepancy could be dependent on the memoranda presentation duration. Malmberg & Nelson (2003) showed that the impact of word frequency on recognition performance (indexed by the hit rate and false detection rate in a change detection task) depended on the presentation duration of the memoranda. The recognition advantage for high frequency words

(i.e., familiar words) compared to low frequency words (i.e., non-familiar words) was only present in conditions with longer presentation duration. The lexical frequency effect disappeared when words were presented for 250ms. In the same vein, Zhou, Mondloch and Emrich (2018) showed that the own-race effect (ORE; the fact that faces from our own ethnicity induce better recognition performance than faces from another ethnicity) also depends on the presentation duration, with smaller ORE induced by shorter presentation duration. The lexical frequency effect and the ORE have been shown to rely on differences in long-term memory representation (Engle, Nations & Cantor, 1990; Morton, 1979; Clutterbuck & Johnston, 2002). More familiar memoranda (high lexical frequency words, own-race faces) have better long-term memory representation than less familiar memoranda (low lexical frequency words, other-race faces). Thus, it could be that the presentation duration we used (500ms) was not enough to create a real difference between the trained and the no-training group.

Another possibility, to explain the absence of training manipulation effect in Experiment 2, could be linked to the difficulty of the task. It may be that, although participants in the trained group effectively had better long-term memory representation of the figure than the no-training group, this difference was not enough to induce a pervasive effect between both groups in the complex-span, but enough to induce a difference in the pre- and post- complex-span tasks. If this is the case, then it would mean that the effect of memoranda familiarity on consolidation would be rather small, which is not congruent with the supposition that familiarity effects stem only from a difference in consolidation speed.

Both possibilities presented above could be compounded by our use of a between-subject experimental design. Between subject manipulation have shown to be less sensitive to experimental manipulation than within-subject manipulation when studying cognitive processes, because within-subject designs let control for inter-personal variance and minimize statistical noise (Charness, Gneezy & Kuhn,

2012). These two factors may have masked the effect of training on performance to the complex-span. Indeed, more familiar memoranda were consolidated around 22ms faster than less familiar memoranda. This is a very small difference, and could have been masked by our use of a between-subject experimental plan.

To our knowledge, our study is the first that manipulated consolidation as a post-mask delay orthogonally to memoranda familiarity. Previous experiments that investigated the hypothesis that more familiar memoranda are consolidated more quickly manipulated consolidation delay as a pre-mask delay and found an effect of this manipulation on indexes of consolidation (Blalock, 2015; Xie & Zhang, 2017). The results from these two studies were congruent with our hypothesis, but their manipulation of consolidation was confounded with encoding processes, limiting the conclusion on the impact of memoranda familiarity on consolidation. Our results more unambiguously show an effect of memoranda familiarity on consolidation, where more familiar memoranda are consolidated more quickly than less familiar memoranda.

5.8 Conclusion

In two experiments, we investigated whether consolidation in working memory was influenced by visuospatial familiarity. Our results were congruent with previous literature on consolidation: memoranda were better consolidated when there was a longer delay between a masked presentation and the beginning of a concurrent task. However, evidence for an impact of visuospatial familiarity on consolidation was more ambiguous.

In Experiment 1, in which familiarity was manipulated within-subject, we found that more familiar memoranda induced shorter Initial-RTs and better recognition performance than less familiar memoranda. This pattern of results is congruent with our general hypothesis, but the conclusion is limited due to the absence of an effect of consolidation delay on recognition performance.

In Experiment 2, in which familiarity was manipulated between-subject and used a recall procedure, we found an effect of consolidation delay on recall performance and Initial-RTs, but evidence against our familiarity manipulation on all measured variables. Due to the absence of an effect of our familiarity manipulation in Experiment 2, it is difficult to interpret the results of this experiment in regard to our hypothesis of an impact of visuospatial familiarity on consolidation. However, it shows that manipulating a post-mask delay indeed manipulated a process involved in the overall maintenance of information in working memory (i.e., consolidation). Thus, if Experiment 1 effectively manipulated this consolidation process, then our study is congruent with the hypothesis that more familiar memoranda are consolidated more quickly than less familiar memoranda, and that this effect is separated from encoding.

However, our results also show that this effect of memoranda familiarity on consolidation is rather small. This makes it difficult to conclude that familiarity effect in working memory is only due to difference in consolidation speed. The literature on the impact of presentation duration on familiarity effect are congruent with the supposition that familiarity affects encoding process. The impact of memoranda familiarity on memory performance thus seems to come from a conjunction of an impact on both encoding and consolidation, at least. Future experiments should aim at investigating the relative weight of consolidation and encoding process in the memoranda familiarity effect.

Chapter 6: Is Guided Refreshing about Attentional Refreshing? Reassessment in the Verbal Domain

Working memory is defined as the part of our cognitive system dedicated to the maintenance and processing of information on short time scales (Baddeley, 1986; Barrouillet & Camos, 2015; Cowan, 1995; Oberauer, 2002; see Logie, Camos & Cowan, 2021, for review). Working memory capacity measured with tasks that require the retention and processing of information are highly predictive of school achievement, fluid intelligence and performance in high-level cognitive tasks (see Conway, Jarrold, Kane, Miyake, & Towse, 2007, for a review). Working memory has also been described as the "hub of human cognition" because of its central role in human cognition and its ubiquitous involvement in everyday tasks (Haberlandt, 1997). Despite a large amount of research dedicated to working memory during the last fifty years, its core functioning is still a matter of debate. In the present study, we aimed at advancing the understanding of working memory maintenance mechanisms.

6.1 Maintenance in working memory

Every theory of working memory has to take into account that people tend to forget things as time goes by. This forgetting has been explained either in terms of temporal decay of the mnemonic traces (e.g., Baddeley & Hitch, 1974; Barrouillet, Bernardin & Camos, 2004; Brown, 1958, for a recent review: Ricker, Vergauwe & Cowan, 2014) or in terms of interference within the content of working memory (e.g., Oberauer & Hein, 2012). Both accounts of forgetting have to postulate the existence of a process that helps the maintenance of the information stored in working memory by counteracting short-term forgetting. Depending on their preferred source of forgetting, models suggest that this could be achieved either by actively maintaining the memory traces or by removing the interfering information. Among the maintenance mechanisms, there is evidence for at least two distinct maintenance mechanisms.

The first proposed mechanism is that verbal rehearsal is a maintenance mechanism specific to phonological information (see Souza & Oberauer, 2018, for a contrary view). It works by actively rehearsing, vocally or sub vocally, the to-be-maintained verbal information. The use of rehearsal seems very pervasive, as it is one of the most commonly self-reported maintenance strategies in working memory tasks (e.g., Belletier et al., in revision; Dunloski & Kane, 2007). It has also been shown multiple times that concurrent articulation is detrimental to the maintenance of verbal information in working memory (Baddeley, 1986; 2012; Belletier et al., in revision; Camos, Lagner, & Barrouillet, 2009).

The second proposed mechanism is attentional refreshing, which uses executive attention to maintain information active in working memory (Barrouillet, et al., 2004; Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; see Camos, Johnson, Loaiza, Portrat, Souza & Vergauwe, 2018, for a review). This mechanism is assumed to work on any domain, as evidence was gathered that impairment of attention impacts similarly verbal, visual and spatial information (Vergauwe, Barrouillet & Camos, 2009, 2010). Contrary to rehearsal which is domain-specific, refreshing is described as a domain-general maintenance mechanism that enhances the activation of encoded information in working memory to keep it accessible for ongoing processing. Accordingly, in working memory tasks that include maintenance of information as well as processing of concurrent information, it has been shown that recall performance depends in a large part on the attentional capture induced by the processing part (see Barrouillet & Camos, 2015, for a review). In other words, when attention is captured by a concurrent task, it is distracted away from maintenance, and recall is impaired. For example, when maintaining letters for subsequent recall, participants exhibited poorer recall performance when the concurrent task on series of digits was a parity rather than a location judgment task, the former being more attentionally demanding (Barrouillet et al., 2007). In another approach, Vergauwe, Camos and Barrouillet (2014) have shown that memory load (i.e., how many items held in working memory at the moment) impacts the

performance on the processing task. In a Brown-Peterson task, when the memory load was increased from 0 to 4, response times to a parity judgement task performed in a series of digits presented after the last letter increased from 744 to 915 ms (Vergauwe et al., 2014). This increase in response latency was interpreted as reflecting the postponement of the parity judgment task while letters were refreshed in a sequential manner, the slope being an estimate of refreshing speed (see also Camos, Mora, Oftinger, Mariz-Elsig, Schneider & Vergauwe, 2019, for similar findings). Together, these results are compatible with the notion that processing and maintenance are trading off attentional resources with each other when performed concurrently.

6.2 Additional evidence for an attentional refreshing mechanism

Most of the evidence supporting the existence of attentional refreshing has been obtained by examining the effect of reduced attention available for maintenance on recall or by measuring the response postponement in concurrent processing task when memory load increased.

Some authors argued that these pieces of evidence are only indirect (Souza, Rerko, & Oberauer, 2015), as previous studies manipulated only the opportunity to use refreshing, and not directly the use of refreshing mechanism. To directly test the role of refreshing on working memory maintenance, three new paradigms have been proposed (Rey, Versace, & Plancher, 2018; Souza et al., 2015; Vergauwe & Langerock, 2017).

Firstly, using response times to memory probes to track the content of the focus of attention over time, Vergauwe and Langerock (2017) found that attention remained focused on the last-presented memory item when there was no time for refreshing to occur in between the list items. Whereas this was not the case when there was time for refreshing to occur in between the list items. This is consistent with the idea that people spontaneously refresh items in the focus of attention when there is time to do so. However, as for the previously mentioned effects related to

refreshing, one could also argue that Vergauwe and Langerock (2017) manipulated the opportunity for refreshing, rather than the use of refreshing *per se*.

Secondly, taking advantage of the preliminary association between tones and visual mask, Rey et al. (2018) showed that recall performance was reduced when these tones were presented in a complex span task when attention was available for refreshing memory items. They considered this finding as a direct evidence of refreshing and suggested that the observed impaired memory resulted from the reactivation of the masks, which disrupts refreshing. However, this finding has been recently called into question in a recent failed replication (Bartsch & Oberauer, in press). Thus, it is not clear whether this paradigm is effectively a valid way to manipulate attentional refreshing.

Finally, the third paradigm, called “guided refreshing”, proposes to test refreshing even more directly and is based on the literature on retro-cueing (e.g., Lepsien & Nobre, 2006). This paradigm is based on the idea that the focus of attention can be redirected to working memory traces by using cues pointing to locations where memory items were presented during the study phase. In Souza et al.’s (2015) seminal article, six colored circles were simultaneously presented on screen on six different locations and participants had to memorize the color of each of them. After the disappearance of the circles, four cues (arrows) successively appeared in the middle of the screen for 500ms each. Every cue pointed to a location where one of the circles was presented. Instructions asked participants to “think of” the colors presented in the cued location when the cue appeared. The authors argued that the act of briefly “thinking of” the color held in working memory would refresh it, such instructions having been implemented previously to induce refreshing (e.g., Camos, Mora & Oberauer, 2011; Johnson, Raye, Mitchell, Greene, Cunningham, & Sanislow, 2005; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Raye, Johnson, Mitchell, Reeder, & Greene, 2002).

During the retention interval, a color could be cued 0, 1 or 2 times. After the last cue disappeared from screen, one of the six locations was highlighted and

participants had to recall the color that had been presented at that location during study, by selecting the color on a color wheel (continuous color reconstruction task). The results showed that the more a color had been cued in the retention interval, the smaller the distance was between the presented color and the participant's response. In other words, higher cueing frequency reduced recall error. Similar findings were observed when colors were sequentially presented (Souza, Vergauwe & Oberauer, 2018). Souza et al. (2015) argued that this cueing frequency effect comes directly from an attentional refreshing manipulation. Their rationale was that "thinking of" an item held in working memory will refresh it, with items that were cued more frequently being refreshed more frequently, hence the increased memory performance.

Since attentional refreshing is defined as a domain-general mechanism, one expects that similar pattern of results would be observed with memory material from other domains, and that the cueing frequency benefit should not be limited to visual materials such as colors. Accordingly, Souza, et al. (2018) tested this effect, using a similar design as Souza et al. (2015), with spatial locations as memoranda, memory responses being reported on a continuous scale as for the colors. Results showed a similar cueing frequency effect in spatial as in visual material, and that the effect is not limited to visual material, but also extends to spatial material.

In the same vein, Souza et al. (2018) tested the cueing frequency effect in verbal material, but this time the pattern of results was more ambiguous. When participants had to do the task with six words sequentially presented as memoranda, the evidence for a cueing frequency effect was absent or minimal in two out of three experiments ($BF_{10} = 1.5, 0.4,$ and 7.85 in Exp. 3a, b, and c, respectively). It was only by pooling the three experiments to enlarge the sample size to 87 participants that there was convincing evidence for an effect of cueing frequency (note that evidence in the visual and spatial domains was obtained with samples between 18 and 30 participants).

However, and contrary to the linear cueing frequency effect reported for visual and spatial materials where differences in recall error were apparent between 0 and 1 cueing as well as between 1 and 2 cueing, the cueing effect for verbal materials was due to the recall improvement between 0 and 1 cueing only, without a clear on recall performance difference between 1 and 2 cues. This latter observation and the weaker effect in verbal compared to visuo-spatial domain cast some doubts on the domain-general nature of the cueing frequency effect.

Moreover, it comes as a surprise that the effect in verbal material was unambiguously present only when participants were not under articulatory suppression (Exp. 3c). In Experiments 3a and 3b, participants had to perform a concurrent articulation throughout the task to impair rehearsal. Such a manipulation should favor the use of refreshing to maintain items in working memory (Camos et al., 2011). In Exp. 3c, participants could have relied on verbal rehearsal to do the task. Cued words would have been more rehearsed than the not-cued, which could explain the cueing benefit detected in this version of the paradigm. Since Exp. 3a and 3b used an articulatory suppression method, it can be expected that these would be the ideal conditions to observe a strong effect of cueing frequency if this effect results from the use of attentional refreshing. Instead, Souza et al. (2018) only found the cueing frequency effect in the experiment *without* articulatory suppression (i.e., the less-than-ideal condition to observe refreshing).

The aim of our research was to reexamine the cueing frequency effect in the verbal domain and to test some possible explanations that could account for the divergence of results reported between visuo-spatial and verbal domains of working memory (see Table 1 for a summary).

Experiments	Material	Articulatory suppression	Recall
Exp. 1	5 letters	With and without AS	Typed recall
Exp. 2	6 words vs 6 colors	With AS	Typed recall for words and reconstruction on color wheel for colors
Exp. 3	4 words vs 4 non-words	With AS	Recognition

Table 1: Summary of the design of the three experiments presented in this study

6.3 The current study

As we just mentioned, impairing rehearsal through articulatory suppression seems to play a moderator role in the cueing frequency effect. However, this was inferred from cross-experiment comparison, and thus, our first experiment aimed to reassess this in a within-subject design. If a strong cueing frequency effect is replicated in the condition without articulatory suppression, and the effect disappears when under articulatory suppression, this would cast doubt on the involvement of refreshing in the guided refreshing paradigm. Blocking rehearsal should force participants to use refreshing. Such a pattern of results would suggest that the cueing frequency benefit in the verbal domain does not stem directly from the attentional refreshing mechanism. Moreover, Experiment 1 tested whether the increase in cueing frequency linearly improves verbal recall performance as reported in visual and spatial material, or whether it differs qualitatively from visual and spatial material such that increasing the number of cueing beyond one has no effect on verbal recall.

Hence, Experiment 1 is a conceptual replication of Souza et al.'s (2018) Experiment 3a with the number of cueing (3 levels: 0, 1 and 2) and the absence or presence of concurrent suppression manipulated within-subject. Contrary to Souza et al. (2018), we presented five letters instead of six words. Foreshadowing the results of Experiment 1, evidence was gathered against a cueing frequency benefit, which contrasted with the findings of Souza et al (2018), and only the manipulation of the

concurrent articulation led to an effect on recall performance. This casts doubts on the notion that the cueing frequency benefit reported by Souza et al (2018) is a robust finding. Experiment 2 aimed at replicating more closely the Souza and colleagues (2018) experiment in the verbal domain.

In a second attempt to replicate Souza et al.'s findings, Experiment 2 compared within-subjects the cueing frequency benefit in verbal and visual materials. Like these authors, we used series of six words as verbal material, each word being presented only once across trials. For the visuospatial domain, we implemented the same design as Souza et al. (2015, 2018), with the presentation of six colored dots and a continuous color reconstruction test using a color wheel. Experiment 2 replicated the cueing frequency benefit in the visual domain as reported by Souza et al (2015, 2018) but, as in our Experiment 1, we found evidence against a cueing frequency effect in the verbal domain.

Finally, our last experiment tested one hypothesis to explain the discrepancy between the verbal and visual domains. This hypothesis is based on the involvement of long-term memory. Words may benefit from stronger and more stable representations in long-term memory than colors, and this may impact the cueing frequency effect. Experiment 3 tested this hypothesis by presenting verbal memory items that either benefit from a previously stored representation in long-term memory (words) or not (non-words). If the discrepancy in findings between verbal and visual domains relies on the involvement of long-term memory, we should observe a cueing frequency benefit for non-words (as previously observed for colors), but not for words (as reported in Experiments 1 and 2).

6.4 Experiment 1

In Experiment 1, we aimed to investigate the implication of rehearsal in the guided refreshing paradigm with verbal material in a within-subject experimental design. In particular, we investigated whether the presence vs. absence of concurrent articulation would modulate the cueing frequency benefit for verbal material. The

same paradigm as Souza et al.'s (2018) Experiment 3 was used, with five letters presented as memoranda instead of six words.

Two alternative hypotheses can be spelled out. According to Souza et al.'s (2018) findings, if the cueing frequency benefit in verbal domain requires the involvement of rehearsal, it should only be observed when participants were not under concurrent articulation, leaving them free to use rehearsal to maintain letters. Alternatively, if the cueing frequency benefit reflects the use of refreshing, it should be observed when participants were under concurrent articulation, favoring the reliance on refreshing to maintain memory items.

6.4.1 Method

Participants and Design

Forty-two students from the University of Fribourg (35 women, mean age 21.1 ± 1.9 years) participated in this experiment. Every participant had normal or corrected to normal vision, signed a written consent form, and received course credit or a cinema ticket for their participation in this experiment. The design was within-subject, participants being involved in eight conditions resulting from the orthogonal crossing of the number of time the probed item was retro-cued (4 levels: baseline (no cues), 0-; 1- and 2-cueing) and the presence or absence of articulatory suppression (2 level: with vs. without). Ethical approval was given by IRB of the University of Fribourg.

Material and Procedure

In this task, participants had to memorize five different letters presented sequentially in boxes presented on screen during a study phase. The memoranda were all consonants of the Latin alphabet aside of W, because it is the only trisyllabic consonant in French. Every letter was used approximately as often across trials. The task started with a fixation cross appearing in the middle of the screen for 500ms. Then, the cross was replaced by five boxes, each placed in a specific place on the screen (Figure 1). The background color of the screen was black, and the boxes were

white rectangles with a thin black border line, each of the same size (4cm X 2cm). Next, 100ms after the appearance of the boxes, the first letter appeared for 900ms in one of the boxes. After a 100ms inter-stimulus interval (ISI) where only the boxes stayed on screen, the next letter appeared in the next box until five letters were presented, one in each box.

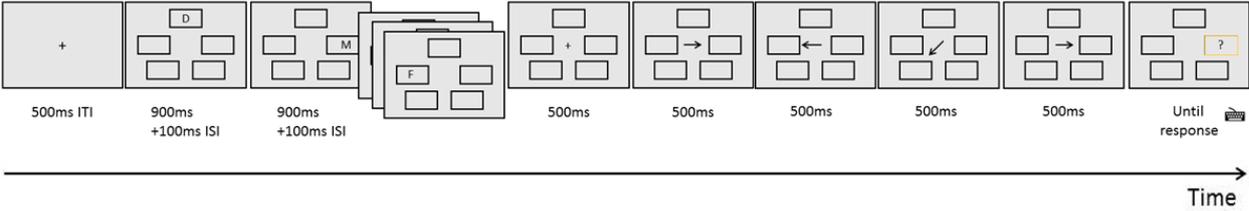


Figure 1: Illustration of a trial in a cued condition in Experiment 1

The first letter was always presented in the top box, and the next letter in the following box going clockwise. The presentation of the last letter was followed by a fixation cross presented in the middle of the screen for 500ms. Then, four arrows were sequentially presented for 500ms each, while the boxes stayed on screen. Each arrow (2cm) appeared at the center of the screen and pointed to one of the five boxes. Participants were instructed to think of the letter presented in the box to which the arrow pointed. After the fourth arrow disappeared, one of the five boxes stayed on screen and its border line was changed to yellow. Participants had to type the letter that was presented in the same box during the study phase. Participants had 10 seconds maximum to type their answer, which corresponds to three times the mean response time of our participants in a pilot study. The letter they typed appeared in the middle of the screen for 320ms. The next trial started as soon as the answer of the participant disappeared of the screen.

Following Souza et al. (2015), an arrow did not point twice in a row to the same box. Following this rule, there were five possible patterns of cueing. If ABCD represent the four letters randomly cued, then, for a given trial, the four arrows could point to four different letters once (A-B-C-D), to two different letters once and to a third twice (three different ways : A-B-C-A ; A-B-A-C ; A-B-C-B) or they could point to

two different letters twice (A-B-A-B). The patterns were the same than those used in Souza et al. (2015) and they were all used approximately the same number of times.

In total, there were 240 experimental trials, half with articulatory suppression (AS) and half without articulatory suppression (No-AS). Both types of trials (AS and No-AS) were presented in separate experimental blocks, the order of which was counterbalanced across participants. In AS trials, from the presentation of the first to-be-remembered letter until they were prompted to recall, participants had to continuously repeat « *ba-bi-boo* » out loud during the whole trial at their own rhythm without pausing with the experimenter monitoring them. This was done to prevent participants from using rehearsal during the task. In this AS condition, « *ba-bi-boo* » was written in the upper half of the screen from the beginning of the trial to the appearance of the yellow box (probe phase) as a reminder. In No-AS trials, participants did not repeat « *ba-bi-boo* » and no indication appeared on screen. The same number of trials was presented for the 0-, 1- and 2- cueing condition (60 trials each: 30 in the AS condition and 30 in the No-AS condition). As in Souza et al. (2018), a baseline condition was added in addition to the cued conditions with 60 trials (30 in the AS condition and 30 in the No-AS condition). In the baseline condition, the cueing phase was replaced by a blank screen for the same duration (i.e., 2 seconds). Aside from the empty cueing phase, the task was exactly the same.

The experiment was divided in four blocks of 60 trials, with two blocks in the AS condition and two blocks in the No-AS condition. There was 15 trials per cueing conditions (baseline, 0-, 1- and 2-cue) in each block. The order of the trials within each block was randomized for each participant. Participants received no indication on the nature of the forthcoming trial. The order of four blocks (2 AS and 2 No-AS) was counterbalanced across participants. A pause occurred between each block, the duration of which was decided by the participants. At the beginning of each block, participants had three training trials, one in baseline, 1- and 2-cue conditions. The experiment was created using Psychopy (Peirce, Gray, Simpson, MacAskill,

Hochenberger, Sogo, Kastman & Lindelov, 2019; version 3.02) and lasted approximately 1 hour in total.

At the end of the experiment, participants had to respond to a surprise questionnaire. This questionnaire was introduced to investigate what participants subjectively did during the experiment. The first question was "Have you been able to follow the instructions and to think about the cued word?". Participants answered on a Likert scale going from 1 "not at all" to 6 "absolutely". If participants scored 3 or less in the first question, they were asked a second question "If you didn't follow the instructions, why is that?" and had a few lines to answer this open question. This questionnaire was done on paper and was used to explore the subjective experience of the participants.

6.4.2 Results and Discussion

Before the analysis, we checked that none of the participants had outlier performance, diverging two MADs (Median Absolute Deviation) from the group median percentage of correct response (median=73.5, MAD=12.0). One participant was excluded on that basis. We also discarded the trials where participants did not respond at all (28 trials in total, i.e., 0.2% of all trials), as we could not infer what participants did exactly in these trials. Since our hypothesis relied on the presence or absence of an effect, we used the Bayesian statistical framework. In this approach, the Bayes factor (BF), which can have a value between 0 and $+\infty$, represents how well a model describes the data compared to another model (e.g., the null model). As such, it can represent how much a model is preferred or not against the null hypothesis. When comparing a given model to the null model, a BF_{10} below 1 means that the null model is preferred, while BF_{10} above 1 means that the tested model is preferred. Due to the continuous nature of BF, a threshold of significance is difficult to evaluate. However, it is widely assumed that any BF_{10} over 3 or under $1/3$ is considered as evidence for or against a model, respectively (Liang & Xiong, 2013). One of the advantages of the Bayesian analysis is that it can provide evidence against the

presence of an effect. Moreover, different statistical models can be directly compared by simply dividing their respective BF_{10} , which estimates how much a model is preferred compared to another one. In addition to the BF of a model, this statistical approach can also give a BF for each factor in the model (called $BF_{\text{exclusion}}$ or $BF_{\text{inclusion}}$), which provides evidence for the absence or the presence of each particular factor, respectively. All analyses have been done using R (R core Team, 2019) and the “BayesFactor” package (Morey, R. & Rouder, J. N. 2018, version 0.9.12-4.2) with the priors set on “medium” for the independent variables effect and set as “nuisance” for the participants ID factor.

We computed the mean recall performance as the number of correct trials divided by the total number of trials per experimental cells. A repeated measure Bayesian ANOVA was performed on recall performance with the cueing conditions (4 levels: baseline, 0-, 1- and 2-cue) and AS (2 levels: AS or No-AS) (Figure 2).

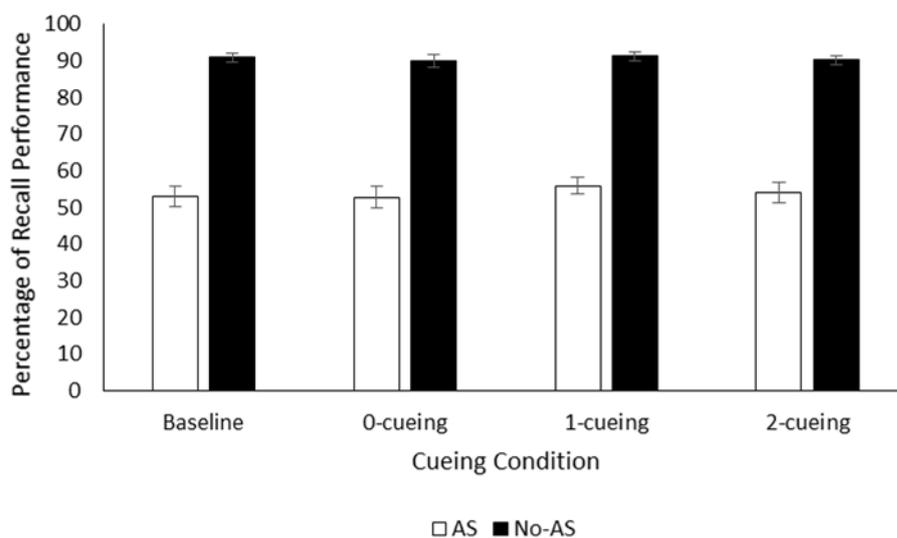


Figure 2: Mean percentage of correct response according to the cueing condition and articulatory suppression in Experiment 1. Errors bars correspond to the standard error of the means.

The participant ID was fixed as random effect, so the actual null model was a simple effect of difference between participants. The best model had only a simple effect of articulatory suppression, $BF_{10} = 4.3 \times 10^{101}$. As can be seen in Figure 2, AS condition yielded worse performance (Baseline: M=53.0, SD=17.5; 0-cueing: M=52.8, SD=18.9; 1-cueing: M=55.9, SD=14.7; 2-cueing: M=54.0, SD=18.0) than No-AS condition

(Baseline: $M=91.0$, $SD=7.5$; 0-cueing: $M=90.0$, $SD=11.1$; 1-cueing: $M=91.3$, $SD=7.8$; 2-cueing: $M=90.2$, $SD=8.2$), as expected. There was evidence against the inclusion of the number of cueing as a predictive factor ($BF_{\text{exclusion}} = 25.6$), which means that the best model explained the data more than 25 times better than a model that also included the main effect of cueing conditions. Similar results were obtained when the baseline condition was discarded from the analysis, thereby providing a more direct test of the cueing frequency effect. The best model included only AS effect, $BF_{10} = 7.9 \times 10^{73}$, and the BF for excluding cueing condition was 7.9. This represented strong evidence against a cueing frequency benefit on recall performance for letters.

Next, we analyzed answers to the questionnaire. We wanted to assess whether the absence of the cueing frequency effect on recall performance depended on the perceived ability to follow the instructions. To that aim, participants were split in two groups according to their answer to the first question (i.e., "Have you been able to follow the instructions and to think about the cued word?"), resulting in a group of 16 participants who answered lower or equal to 3 (named "low group"; mean=2.3, $SD=0.8$), and a group of 25 participants who answered higher than 3 (named "high group"; mean=4.7, $SD=0.6$). A Bayesian ANOVA with the same design (i.e., with the 4 levels of cueing and the AS manipulation) as previously was performed separately on each group. For both groups, the best model had only a main effect of the articulatory suppression condition ($BF_{10}=6.1 \times 10^{41}$ in the low group, and $BF_{10}=1.3 \times 10^{60}$ in the high group) with strong evidence against a main effect of the cueing condition ($BF_{\text{exclusion}} \text{ cueing frequency}=10.1$ and 20.7 , in the low and high group respectively). Similar findings were obtained when the baseline condition was discarded ($BF_{\text{exclusion}} \text{ cueing frequency}=4.2$ and 7.8 , in the low and high group respectively). In the low group, six participants reported that the arrows were too fast to follow and seven said that repeating *ba-bi-boo* made the overall task very hard. The five other participants just said that the task was difficult. Participants in the high group did not have to respond as to why they felt like they succeeded to follow the task rules.

To summarize, Experiment 1 provided evidence against the cueing frequency benefit on letters in the guided refreshing paradigm. Contrary to Souza et al. (2018) who reported an effect when rehearsal was available and no effect when it was impaired, evidence in Experiment 1 was always against the cueing frequency benefit regardless of whether participants could use rehearsal ($BF_{\text{exclusion}} = 18.8$ and 8.7 , with and without the inclusion of the baseline, respectively) or not ($BF_{\text{exclusion}} = 7.5$ and 3.5 , with and without the inclusion of the baseline, respectively). However, as expected, our participants had very good recall performance when they could use rehearsal and thus, it could be argued that a ceiling effect in the condition without articulatory suppression could mask any cueing benefit. Hence, Experiment 2 attempted to replicate Souza et al.'s (2018) findings by following more closely these authors' design in the verbal domain. Moreover, we also included a replication in the visual domain, allowing us to proceed to a within-subjects comparison of the cueing frequency benefit in verbal and visual domains.

6.5 Experiment 2

To replicate more closely the design of Souza et al. (2018), Experiment 2 used words as verbal memory items instead of letters and the number of memory items was increased to six instead of five. Each word was used only once across trials. For the visual domain, we used the available Souza et al. (2015, 2018) script.

6.5.1 Method

Participants and Design

Forty students (34 women, mean age 21.7 ± 1.2 years) from the University of Fribourg participated in the experiment. Every participants had normal or corrected to normal vision and no color blindness and signed a form of consent. Each participant received course credits or cinema ticket for their participation. None of them participated in Experiment 1, but they all participated to the two parts of this experiment (within-subject design). The experiment was divided in two parts, the

order of which was counterbalanced across participants. One part was the verbal version of the task with words as memoranda and the other was the visuospatial version of the task, which used colored dots as memoranda. Ethical approval was given by IRB of the University of Fribourg.

Material and Procedure

Experiment 2 was divided in two parts. Both parts used a guided refreshing paradigm but in a different domain, verbal and visuospatial with either words or colors as memoranda. For the verbal domain, the same paradigm was used as in Experiment 1 with six words instead of five letters and all trials were performed under articulatory suppression to minimize rehearsal and favor the use of refreshing. Contrary to Experiment 1, participants could correct their answer using backspace. Once they were satisfied with their answer, they pressed the "enter" key and the next trial started. As in Souza et al. (2018), each word was presented only once in the experiment. There was a total of 96 trials, 24 in each cueing condition as well as in the baseline condition. Trial order was randomized for each participant.

Six hundred French monosyllabic words were selected as memoranda. They were selected in a database of 1493 French words (Bonin, Méot, Ferrand & Roux, 2011). To ensure that the words were well-known by the participants, we kept the 600 most frequent monosyllabic words (mean lexical frequency in books = 55.8; SD = 134.9, ranging from 0.14 to 1140 occurrence per million).

For the visuospatial domain, we had access to the script of Souza et al. (2015, 2018), and we implemented it with no change. At the beginning of the trial, six colored circles appeared in the middle of the screen (Figure 3). Participants had to memorize the colors and where they were placed. After a 500ms interval, four arrows pointed randomly to where circles were presented, and participants were tasked to "think of" the colors that were presented in the circle to which an arrow pointed. After the last arrow disappeared and a 500ms interval, an empty circle appeared where one of the colored circles was presented at the beginning of the trial. Participants had to

click using the computer mouse on a color wheel what color they thought was presented in the highlighted circle at the beginning of the trial. We used the distance in angular degree between the probed color and the participant’s answer as dependent variable (hereafter: absolute recall error).

As in the verbal domain, to hinder the use of the rehearsal, participants repeated *ba-bi-boo* out loud in each trial from ready signal to recall prompt. Colors were sampled from a continuous CIE L*a*b color model (L = 70, a = 20, b = 38, and radius = 60) with 360 values evenly distributed along a color wheel (Zhang & Luck, 2008). They were randomly selected with the constraint that all six colors within a trial were at a minimum distance of 20° of each other on the color wheel. There was a total of 150 trials, 50 in each cueing condition (0-, 1- and 2- cueing).

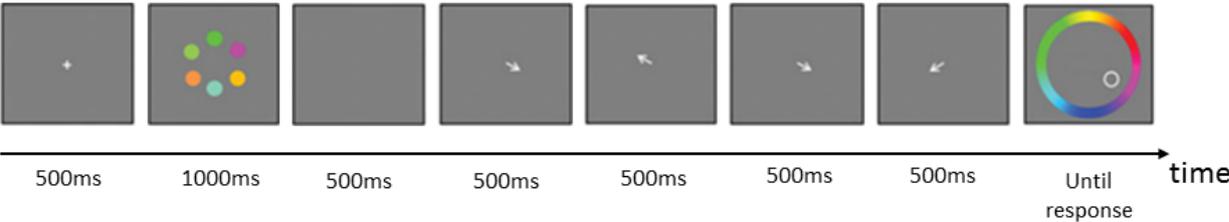


Figure 3: Illustration of a trial for a cued condition with colors in Experiment 2.

All participants did both parts (i.e., verbal and visuospatial domain), the order of which was counterbalanced across participants. Like in Experiment 1, every participant answered a post-experiment questionnaire for each part of the experiment. The questionnaire was prompted after the participant ended both experimental parts. The first two questions were “did you try to think of the cued word/color?” and “did you succeed in thinking of the word/color when its box/circle was cued?”. Participants answered these questions on a Likert scale going from “1” (not at all) to “6” (absolutely). The third question was an open one. Participants had to describe what obstructed them from thinking of the cued words/colors. Finally, they had to describe what kind of strategies they used in the task. The word part has been programmed using Psychopy (Peirce et al, 2019) and the color part has been

programmed using Matlab (ver.2017b) and the Psychtoolbox module (Brainard, 1997).

6.5.2 Results and Discussion

Performance in the verbal and visual domains was analyzed separately as different dependent variables were used (percentage correct for words, absolute recall error for colors). In both tasks, no participants had performance diverging two MADs or more from the group median. First, as in Experiment 1, a repeated measure Bayesian ANOVA was performed on the percentage of correct word recall with the cueing conditions (4 levels: baseline, 0-, 1-, and 2-cueing) as within-subject variable (Figure 4A). The model yielded a BF of 3.3 against the inclusion of the number of cueing as a predictive factor (Baseline: M=43.3, SD=16.4; 0-cueing: M=39.0, SD=16.2; 1-cueing: M=39.4, SD=14.6; 2-cueing: M=42.0, SD=15.9). As in Experiment 1, we did the same analysis without the baseline condition, and this analysis yielded the same pattern of results, with evidence against the effect of the number of cueing, $BF_{01} = 5.1$.

For the visual domain, we did the analysis on the mean absolute error (in degree) in the continuous color reconstruction task yielded a BF of 10.9 for the effect of the cueing condition (0-cueing: M=71.0, SD=53.9; 1-cueing: M=69.3, SD=54.6; 2-cueing: M=64.1, SD=53.3).

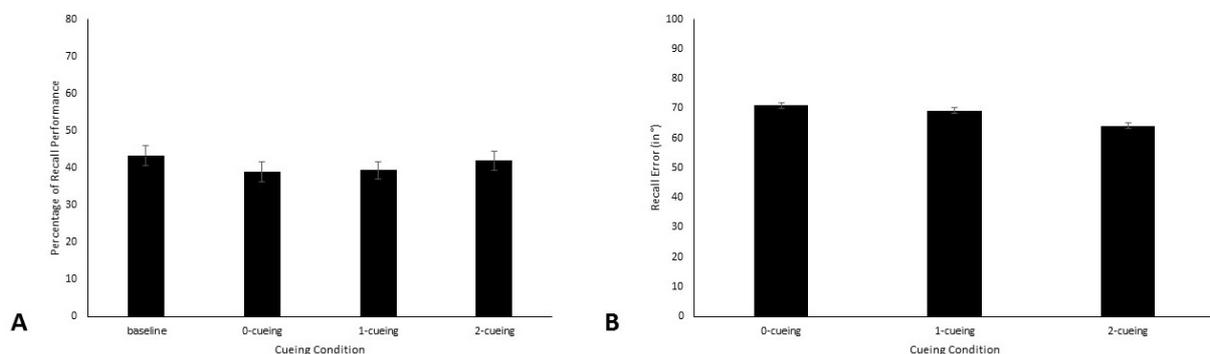


Figure 4: (A) mean percentage of word recall performance and (B) mean recall error (in degree) for colors according to the cueing conditions in Experiment 2. Error bars represent the standard error of the means.

Thus, in the same sample, we replicated the presence of the cueing frequency benefit in the visual domain but could not gather evidence for an effect of the same variable in the verbal domain, even though our experiment followed closely the paradigm from Souza and colleagues (2018).

The analysis of the post-experiment questionnaire showed that participants tried to follow the instructions related to the cues in the word and the color task, ($M=3.7$, $SD=1.6$, for both tasks for the first question), but participants did not feel like they were succeeding in following these instructions ($M=2.7$, $SD = 1.4$ and $M=2.8$, $SD= 1.1$, for the second question of the word and color task, respectively). To the open question asking participants why they did not succeed to follow the instructions, more than three quarters of the participants reported that the cues appeared too quickly on screen to think about all the retro-cued items in both part of the experiment.

As in Experiment 1, we reanalyzed our data in regard to the answer to the post-experiment questionnaire. First, we analyzed the answers to the word part. A Bayesian ANOVA with a similar design as the previous one was performed on the two groups of participants segregated on their score on the first question "did you try to think of the cued word?". In the high group (26 participants who scored above 3), the BF for the presence of the cueing benefit was 1.04 when the baseline condition was included, meaning that the model including the number of cueing factor did not explain the data better than the null model. The same analysis without the baseline condition yielded a BF against the presence of the number of cueing frequency effect of 2.7, which can be considered weak evidence against the presence of the effect. In the low group (14 participants who scored below 3), the best model was the null, with a BF of 9.8 and 5.7 against the presence of the cue benefit when the baseline condition was included and discarded, respectively. We similarly segregated participants in two groups, based on their answer to the second question "Did you succeed in thinking of the word when its box was cued?". One group (10 participants, high group) included participants who answered more than 3, and the other group

(30 participants, low group) participants who answered 3 or less. In both groups, evidence favored the absence of the effect of the cueing condition, with a BF against the effect of 6.6 and 3.1 for the high and low group, respectively. The exclusion of the baseline condition did not change the pattern of results, with evidence against the effect of the number of cueing (BF_{01} of 4.4 for both groups).

The same analyses were performed on the color task. First, participants were segregated into two groups based on their answer to the first question. The analysis on the high group (22 participants scoring above 3) showed an effect of the number of cueing, with a BF of 17.4 for the presence of an effect. Interestingly, the same analysis on the low group (18 participants scored 3 or below) showed evidence for the absence of an effect of the number of cueing, with a BF against the effect of 5.1. For the second question, in the high group (11 participants), evidence for the effect of the number of cueing was ambiguous, with a BF against the effect of 1.5. In the low group (29 participants), the effect was more clearly present, with a BF for the presence of an effect of the cueing condition of 3.1.

In response to the third question "If you couldn't think of the cued item, what stopped you?", 21 out of 40 participants reported for both tasks that the cues appeared too quickly on screen. It was the most reported reason, with the only two other reasons mentioned for the two tasks by more than one participant being: distracted by the cues themselves (4 participants) and distracted by the articulatory suppression (3 participants). It should be noted that participants reported the same difficulties to think of the cued items in the two tasks (i.e., for words and colors). Finally, we assessed the strategies reported by our participants in the last question of the questionnaire. In the verbal part, the most reported strategy was elaboration (25 participants out of 40), where participants bundle together items by creating a story or a sentence containing the presented items. The second most reported strategy was rehearsal (10 out of 40 participants). Although they were under articulatory suppression, some participants expressed they were able to nevertheless rehearse the memory words. In the visual part, the most reported strategy was categorization,

where participants put the different colors into more broad categories, like cold or hot colors, and use these labels to help them remembering the colors (15 out of 40 participants). The second most used strategy was some kind of binding strategy, where colors that were close in the color space were bundled together (13 out of 40).

Thus, like in Experiment 1 and although we used the same type of material and the same number of memory items as Souza et al. (2018), we still found evidence against a cueing frequency benefit in the verbal domain in Experiment 2 (Figure 4B). However, although our participants' performance was a bit lower than in previous experiments, we did replicate Souza et al.'s (2015, 2018) finding in the visual domain. How can the divergence between the two domains be understood? First, a difference between the two types of material relies on the fact that verbal items (letters or words) are easily nameable while colors were chosen to be difficult to be named. One would think that the involvement of some maintenance mechanisms specific to verbal information, such as rehearsal, could account for this discrepancy. However, we took care in Experiment 2 to minimize rehearsal and labeling by the presence of a concurrent articulation in both the word and color tasks. Moreover, in both tasks, similar numbers of participants reported the use of verbal strategies (i.e., 10 in word and 15 in color tasks). Hence, it is very difficult to consider verbal strategies as the source of differences between the two tasks.

A second difference between the verbal and visual material could be related to the degree of involvement of the long-term memory. Both the letters in Experiment 1, and the words in Experiment 2, benefit from well-established representations in long-term memory. As a consequence, maintenance and recall of verbal material can benefit from the support of long-term memory more easily than colors, which need an additional labeling step before a verbal long-term memory representation can be used to maintain it. Accordingly, the majority of participants reported the use of elaboration, a maintenance strategy relying on long-term memory, for the word task, something that was not reported for the color task. Also, verbal representations are much more categorical in nature than visual representations, which are more

continuous, all the more for colors. In the verbal domain, memory items (or some of them) can be stored in secondary memory, which reduces the load in working memory. At recall, items stored in secondary memory are retrieved by a search process, using other cues than the word-location bindings (Unsworth & Engle, 2007). Verbal memory traces can also benefit from reconstruction based on knowledge stored in long-term memory by redintegration (Hulme, Roodenrys, Schweickert, Brown, Martin & Stuart, 1997; Lewandowsky & Farrell, 2000). Experiment 3 examined this proposal by manipulating the implication of long-term memory for the verbal memoranda in guided refreshing paradigm. Therefore, we manipulated the lexicality of the verbal items to be maintained. We reasoned that, if support from long-term memory hinders the cueing frequency benefit for letters and words, this should not be the case for items that are less well represented in long-term memory. Therefore, non-words, which have less stable long-term memory representations, should benefit more from the number of cueing, as it is the case for colors.

6.6 Experiment 3

In Experiments 1 and 2, we found evidence against an effect of the number of cueing on recall performance of letters or words. However, in line with Souza et al. (2015, 2018), Experiment 2 found evidence for the cueing frequency benefit in the visual domain when colors had to be maintained. One hypothesis to explain this discrepancy could be linked to the involvement of long-term memory. To test this hypothesis, Experiment 3 aimed at varying the reliance on long-term memory in the verbal domain by manipulating the lexicality of the memory items. Therefore, we compared words to non-words, both being verbal material but the latter minimizing the involvement of long-term memory. We also wanted to use a more fine-grained measure to favor the detection of the cueing frequency benefit. Hence, we used a recognition task for which response times to the presented items can be measured as well as the rate of correct responses.

6.6.1 Method

Participants and Design

Thirty students from the University of Geneva (25 women, mean age 20.7 ± 2.5 years) participated in the experiment. Every participant had normal or corrected to normal vision, signed a consent form before participating in the experiment, and received course credit for their participation. The number of time the probed item was cued (4 levels: baseline, 0-, 1- and 2-cueing) and the lexicality of the memory items (2 levels: Words or Non-Words) were orthogonally manipulated in a within-subject design. Ethical approval was given by IRB of the University of Fribourg.

Material and procedure

As in Experiment 1, memory items were presented in boxes, here two lines of two boxes to present four items. The size of each box was 5cm x 4cm and each box had a thin, black border line. Each item was presented in the center of one of these boxes, in Courier New font, 24 points.

Memory items were either words or non-words. One hundred and twenty words were selected from a set of 327 monosyllabic singular French nouns with a Consonant-Vowel-Consonant phonological structure and with mean frequency of 59 (SD = 147; range from .27 to 861 frequency of occurrence in books) extracted from Lexique3 (New, Brysbaert, Veronis & Pallier, 2007). Based on these words, 120 non-words were created using Lexique Toolbox by substituting one phoneme of a word by another phoneme randomly selected by the software. Iterations resulting in unpronounceable non-words, pseudo-homophones and pseudo-words with more than three phonemes were discarded and the script was re-ran. This material was pre-tested by Camos et al. (2019;) and induced a strong lexicality effect in a simple span task, with better recall performance for words (49%, SD = 21) than non-words (22%, SD = 11), $BF_{10} = 5.7 \times 10^5$. Words and non-words were segregated in two lists, L1 and L2. Half of the participants were presented with words of L1 and non-words of L2,

and the other half with the rest of the material. On each trial, four memoranda were chosen randomly without replacement from the list given to the participants. Across the experiment, each item was presented 8 times as memorandum.

Each series began with a fixation cross that was centrally displayed on screen for 500ms. The next screen added the four boxes and, in the left box of the upper row, the first to-be-remembered item was presented for 900ms, followed by a screen showing the four empty boxes for 100ms. Next, the four boxes were again presented, this time with the second to-be-remembered item presented for 900 ms in the right box of the upper row, again followed by a screen showing four empty boxes for 100ms. This continued for the third and fourth items (with the third item shown in the left box of the lower row, and the fourth item in the right box of the lower row). The presentation of the fourth item was followed by a 2.5 sec-delay (500ms of pre-cueing blank plus 2000ms of cueing interval). During the cueing interval, four arrows appeared sequentially on screen for 500ms each, as in the previous experiments. After the retention interval, an item appeared in one of the box, and participants had to decide if it was the same item than the one presented at the beginning in this box. On half of the trials, it was the item presented in this box, and on the other half it was an item randomly selected from the other list than the one assigned to the participant, and hence never studied. Participants answered using the keys "b" and "c" to answer "yes" or "no" respectively. They had no maximum time limit to answer, but they were instructed to respond as fast and as correctly as possible. Once they answered, a feedback was given by showing "correct" or "false" in green or red, respectively, in the box the probed item was in for 500ms. The next trial started when the feedback disappeared from the screen. The correctness of the answer as well as the reaction time were recorded. To ensure that the effect between the words and the non-words was not linked to the use of verbal rehearsal, participants had to repeat "ba-ba-ba" from the beginning of the trial until the response phase (but not during response). We recorded the response time and the correctness of the answers.

There were 32 trials for each experimental condition crossing the number of cueing (3 levels: 0, 1- and 2-cueing) and the lexicality of the items (2 levels: words vs. non-words). In addition to these 192 trials with cue, there were 48 baseline trials randomly inserted in between the other trials. These trials had the exact same structure than the others, with the difference that no cue was presented during the retention interval. Of these 48 baseline trials, 24 used words and the other half used non-word as memoranda. The order in which the trials appeared was randomly selected for each participant.

Before the beginning of the experimental trials, each participant did 8 training trials. First, they had a trial without cues (baseline) and articulatory suppression, then a trial with cues but no articulatory suppression, then 6 trials with cues and articulatory suppression. Words and non-words trials were used for the training trials. There was a pause every 60 experimental trials.

6.6.2 Results and Discussion

Three participants had their median percentage of correct responses below 2 MAD of the group median and were excluded from the analysis. First, we computed the percentage of correct responses to assess recognition performance. A Bayesian ANOVA was performed on this score with the lexicality and the cueing condition as within-subject variables (Figure 5A). The best model had only a main effect of lexicality, $BF_{10} = 21.2$. The non-words yielded worse recall (baseline: $M=81.5$, $SD=10.0$; 0-cueing: $M=81.0$, $SD=7.7$; 1-cueing: $M=80.7$, $SD=7.1$; 2-cueing: $M=82.1$, $SD=7.1$) than words (baseline: $M=86.0$, $SD=7.3$; 0-cueing: $M=83.3$, $SD=8.3$; 1-cueing: $M=84.1$, $SD=8.0$; 2-cueing: $M=84.4$, $SD=8.2$). The BF for excluding the effect of cueing condition as a predictive factor was 22.9, and thus there was strong evidence against a cueing frequency benefit. There was also evidence against the interaction between the two variables, with $BF_{\text{exclusion}} = 14.1$. Removing the baseline condition did not change the pattern of results, with the best model still being a simple effect of

lexicity ($BF_{10}=4.0$) and evidence against the effect of the cueing condition ($BF_{\text{exclusion}}=11.3$) and the interaction ($BF_{\text{exclusion}}=8.7$).

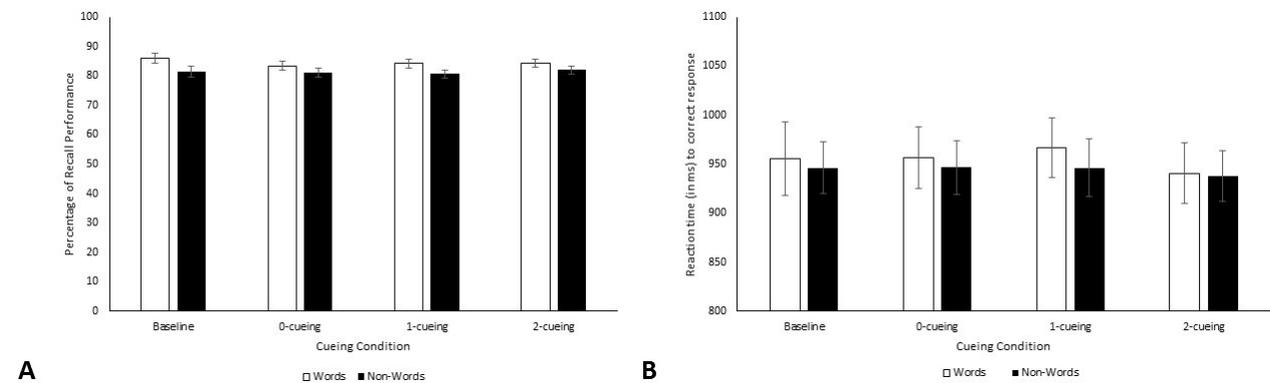


Figure 5: (A) Mean recognition performance (in percentages) and (B) mean response time (in ms) of correct trials according to the cueing condition and the lexicity of the memory item in Experiment 3. Error bars correspond to the standard error of the mean.

For the analysis of the response times, we discarded every trial where participants failed to answer correctly and computed the mean response time per participant and per experimental condition (Figure 5B). A Bayesian ANOVA on the mean response times with the cueing condition and the lexicity condition as within-subject variables was performed. The best model was the null model, with BF against the number of cueing at 23.2, against the lexicity effect at 4.6 and against the interaction at 17.2 (Words: baseline: $M=955$, $SD=197$; 0-cue: $M=957$, $SD=163$; 1-cue: $M=966$, $SD=157$; 2-cue: $M=941$, $SD=162$; Non-words: baseline: $M=96$, $SD=136$; 0-cue: $M=947$, $SD=142$; 1-cue: $M=946$, $SD=154$; 2-cue: $M=938$, $SD=137$). Removing the baseline condition did not change the pattern of results, with still the null model as the best model (Number of cueing: $BF_{\text{exclusion}}=6.7$; lexicity: $BF_{\text{exclusion}}=3.5$; interaction: $BF_{\text{exclusion}}=7.7$).

Thus, in Experiment 3, we could again not replicate the results from Souza and colleagues for verbal material, examining the cueing frequency benefit on recognition performance as well as on response times. In a last attempt to detect the cueing frequency effect in verbal domain, we concatenated the data from our three experiments as Souza and colleague (2018) did for their Experiment 3a, 3b and 3c.

6.7 Cross-Experiment Analysis

Like Souza et al. (2018), we pooled the data of our experimental conditions that included verbal material (Table 1: all conditions from Experiments 1 and 3, and the word task in Experiment 2), and did a new Bayesian ANOVA on recall performance with the number of cueing (3 levels: 0-, 1- or 2-cueing) as within-subject factor and Experiment as between-subject factor. The best model had only a main effect of Experiment ($BF_{10} = 1.8 \times 10^{27}$). The BF for the exclusion of the number of cueing was 5.9 and against the interaction was 18.2. However, it remained possible that there were too many differences between experiments (presence vs. absence of articulatory suppression in Experiment 1, words vs non-words in Experiment 3). Hence, we did the same analysis, but only with the AS condition in Experiment 1, and the words for Experiments 2 and 3. This yielded the same pattern of results as previously, with the best model being the one with a simple effect of Experiment ($BF_{10} = 1.3 \times 10^{20}$) and a BF against the inclusion of the cueing condition of 9.4 and against the interaction of 26.6.

In further analyses, we examined the difference between each cueing condition against the baseline condition. We found weak evidence for worse performance in the 0-cueing compared to the baseline condition ($BF_{inclusion} = 2.0$), weak evidence against a difference between the 1-cueing and baseline condition, ($BF_{exclusion} = 2.7$), and moderate evidence for an absence of difference between the 2-cueing and baseline condition ($BF_{exclusion} = 6.1$).

In their study, Souza and colleagues (2018) found that their effect on verbal material was driven by the difference between the 0- and 1-cueing conditions. For this reason, we also analyzed the difference between adjacent levels of the cueing conditions. To this aim, we did the same Bayesian ANOVA as before, but with only the levels of interest of the cueing conditions. This always yielded the same pattern of results: the best model was the one that only included a simple effect of Experiment, with evidence against an effect of the cueing condition ($BF_{exclusion} = 3.7$ for 0-vs. 1-cueing; $BF_{exclusion} = 5.3$ for 1- vs. 2-cueing). To summarize, we systematically gathered

evidence against the cueing frequency benefit in verbal material across all our experiments.

6.8 General Discussion

In this series of experiments, we investigated the guided refreshing paradigm in the verbal domain. Previous research using this paradigm found a robust effect of the number of cueing in the visual domain, but the effect was much more ambiguous with verbal memoranda. Because this paradigm is assumed to investigate attentional refreshing in working memory, and attentional refreshing is defined as a domain-general mechanism, the ambiguous effect in the verbal domain casts doubt on the idea that the paradigm is indeed assessing the use of refreshing.

In the first experiment, we used the guided refreshing paradigm with letters as memoranda and manipulated the presence (AS) or absence (No-AS) of articulatory suppression. Articulatory suppression minimizes the opportunity to use rehearsal as a maintenance strategy. As a consequence, participants should rely more on attentional maintenance mechanism in the condition with articulatory suppression. The cueing frequency effect should be particularly present in this condition, if it relies on attentional refreshing. However, our results showed evidence for the absence of an effect of the number of cueing, both in the AS and No-AS conditions. Interestingly, the evidence for the absence of the effect of the number of cueing was present for participants who subjectively felt that they followed the instructions as well as for participants who felt that they could not.

In the second experiment, we switched to words as memory items to increase the difficulty of the task and to follow more closely the design of Souza et al. (2018). Moreover, our participants also did Souza et al.'s (2015, 2018) visual version of the task to ensure that any difference in results between our study and the previous studies by Souza and colleagues is not related to differences in the participant sample. Again, no effect of the number of cueing on recall performance was detected in the verbal domain, even though we replicated in the same sample the cueing

frequency benefit with colors as memoranda that Souza et al. (2015, 2018) observed. The response to a post-experiment questionnaire investigating the subjective feelings of the participants showed that even when participants tried to follow the instructions and felt that they were succeeding, there was still evidence for the absence of an effect in the verbal domain. On the contrary, in the visual task, only participants that effectively tried to follow the instructions showed an effect of the number of cueing. Participants that subjectively felt that they could not follow the instructions showed evidence for the absence of an effect of the number of cueing in the visual domain. To our knowledge, our study is the first that explores the subjective feelings of participants in this paradigm, and the latter finding brings some nuance to the existence of the cueing frequency benefit in the visual domain.

Finally, in the last experiment of our study, we slightly modified the paradigm to get more fine-grained measures, collecting response times in a recognition procedure. We also manipulated the lexicality of the to-be-maintained items, because reliance on long-term memory could have hindered the effect of the number of cueing for words and letters, whereas this should not be the case for non-words, which can benefit less from long-term memory knowledge. Once more, we found evidence against the effect of the cueing conditions on memory performance of verbal items, for both recognition accuracy and response times. Even when we pooled the data from all our experiments together (109 participants in total), the evidence still favored the absence of an effect of the number of cueing in verbal material.

To summarize, we did not replicate the results from Souza et al. (2018) in the verbal domain, consistently providing evidence against it instead. Moreover, we showed that the cueing frequency benefit in the visual domain is not ubiquitous among participants, who are able to subjectively evaluate how effective the cueing was for their performance. In the following, we enlighten some aspects of the findings of Souza et al (2015, 2018) that nuance their conclusion on the existence of a cueing frequency benefit in verbal working memory, and that are, in fact, in line with

the present study. Finally, we propose some accounts for the different patterns observed in the visual and verbal domains of working memory.

6.8.1 Is there a real divergence of results in the verbal domain?

Overall, our study provided evidence against the effect of the number of cueing in the verbal domain, which contradicts the conclusion of Souza and colleagues (2018). However, a more detailed examination of the findings of Souza et al. (2018) enlightens this apparent inconsistency, because three aspects of their findings tone down their conclusion. First, when Souza et al. (2018) analyzed across their three experiments the effect between adjacent levels of the number of cueing separately, evidence was gathered for an effect of cueing between the 0- and 1-cueing conditions ($BF_{inclusion} = 5.5$), but the evidence was ambiguously against the effect between the 1- and the 2-cueing conditions ($BF_{inclusion} = 0.5$).

This pattern is rather striking, because the idea behind the guided refreshing paradigm is that recall performance should monotonically increase with the number of cueing. Indeed, the paradigm was created on the basis that varying the number of cues directly manipulates refreshing activities, and that the number of cueing directly correlates with the number of refreshing steps. If the difference is only between items that were never cued versus items that were cued once or twice, the pattern of results resembles more closely a simple retro-cue effect than a cumulative one, even though the cues in the guided refreshing paradigm are much less predictive of the item to be tested than in the retro-cue paradigm. Indeed, it has been shown that using a single retro-cue in a working memory task can boost performance for the cued item by retroactively focusing attention on the cued item (see Souza & Oberauer, 2016, for a review). Usually, this benefit comes at the cost for non-cued items. However, a "simple" effect of cued vs. non-cued items on recall performance is not the same as a cumulative retro-cue effect, where several cues can point to a single item, and each cue in itself boosts performance for this item. Thus, the conclusion of Souza et al. (2018) should be nuanced, because, contrary to the visual domain in which they

reported a monotonically increasing recall performance, performance was only improved by a single cueing and there was no evidence for a cumulative cueing effect for verbal material. This qualitative difference in the effect of cueing frequency on memory performance between the visual and the verbal domains of working memory casts doubt on the idea that the same mechanism is at play in the two modalities. We return to this point in the following section.

Second, as we previously discussed in our introduction (chapter 6.2), Souza and colleagues found convincing evidence for an effect of the number of cueing in verbal domain only when participants were not under articulatory suppression (Exp. 3c). In the two other experiments under articulatory suppression (Exp. 3a and 3b), their evidence was ambiguous regarding an effect of the number of cueing. This is surprising, because articulatory suppression is thought to minimize the opportunity to use verbal rehearsal as a maintenance strategy and thus to maximize the use of attentional refreshing (see Camos et al., 2018). If this is true, the number of cueing effect should have been more pronounced under articulatory suppression, which is not what Souza and colleagues observed. The fact that they only observed convincing evidence for the number of cueing effect on verbal materials when participants were able to freely use verbal rehearsal raises the possibility that the observed effect is related to the use of verbal rehearsal. Participants could have overtly or covertly rehearsed the cued words more than the non-cued ones, and this would give them a recall benefit compared to the non-cued words. However, we directly tested the moderating effect of the presence or absence of articulatory suppression on the cueing effect, but it did not change the evidence against the effect of number of cueing.

Finally, another striking feature in Souza et al. (2018)'s data is the difference between the baseline condition and the cueing conditions. Surprisingly, when the data were pooled across three experiments, results favored the absence of an effect between each cueing condition and the baseline, with tentative evidence for better performance in the baseline compared to the 2-cueing condition. In our experiments,

the evidence was against a difference between the baseline and the 1- and 2-cueing conditions, and data favored (although ambiguously) worse recall for the 0-cueing condition compared to the baseline. This is all the more surprising as one may think that the effect of the number of cueing should give a benefit compared to the baseline, at least for the 1- and 2- cueing conditions. Indeed, items that are refreshed can be thought as prioritized against the other items maintained in working memory, and thus better recalled (Bartsch, Singmann & Oberauer, 2018; Morey, Cowan, Morey & Rouder, 2011; Rhodes et al., 2019; Sperling & Mechner, 1978). In the baseline condition, no specific item could be prioritized for maintenance compared to each other. In the 1- or 2-cueing condition, insofar as retro-cues actually prompt participants to refresh the cued items, they should be prioritized compared to non-cued items. This prioritization should give a recall advantage for any cued condition compared to the baseline, but this is not the case in our data as well as in Souza et al. (2018).

One possibility is that, in the baseline condition, verbal memory items are spontaneously refreshed in the most optimal way, making it extremely hard to find any evidence for a beneficial effect of guided refreshing in the verbal domain (see Bartsch et al. 2018, for a similar proposal). In line with this idea, Vergauwe (2018) observed that verbal recall performance was not influenced by the varying nature of guided refreshing schedules (e.g., refreshing of list items in forward order starting with memory item 1 vs. refreshing list items in a random order) and, more importantly, that all guided refreshing schedules resulted in poorer recall performance relative to the baseline condition.

To conclude, the absence of a clear difference between the baseline and the cueing conditions, as well as the difference being present only between the 0- and 1-cueing condition in Souza et al.'s (2018) data complete the doubts mentioned in introduction concerning the cueing benefit effect under articulation suppression. It thus appears that, overall, Souza et al.'s (2018) data are not that different from our pattern of results after all, and they do not seem to support the conclusion that the

same mechanism is at play in the verbal and the visual domains. This casts doubt on the domain-generality of the cueing frequency benefit, and thus on the fact that the guided refreshing paradigm provides a direct test of attentional refreshing as advocated by Souza et al. (2015, 2018). Furthermore, the question remains as to why the effect is different between the two domains of working memory. At least three accounts can be outlined.

6.8.2 Potential accounts for a divergence between the verbal and the visual domains

One potential difference between the two domains could be the way participants handled the instructions. Our analysis of the post-experiment questionnaire in Experiment 2 showed that participants tended to think they were similarly engaged in the verbal and visual tasks. However, in the visual task, there was evidence for the effect of the number of cueing only in the group that subjectively felt that they could follow the instructions, with evidence of the absence of an effect in the other group. This was not the case in the verbal task, in which consistent evidence was gathered against an effect of the number of cueing, regardless of whether participants felt that they could follow the instructions or not. This reinforces the idea that different processes are involved in the two domains. Moreover, participants' reports in the visual task can result from two distinct mechanisms.

Indeed, participants could interpret the instructions as requiring that a particular maintenance strategy is to be implemented, and only those participants who succeeded in implementing the strategy showed the frequency cueing benefit. This would suggest that the maintenance strategies are consciously controlled and is in line with Souza et al.'s (2014) proposal that attentional refreshing is a strategic process operating at long timescales (several hundred of milliseconds). This view contrasts with the idea that attentional refreshing works on shorter time scale of 50-100ms (Vergauwe et al., 2014) and thus occurs largely outside of our awareness. Although the two accounts are not necessarily exclusive, because both kinds of

processes (a long and a short refreshing) could exist in parallel to boost recall performance (see Camos et al., 2018, for a similar proposal), our results are at odds with both accounts. If attentional refreshing operates automatically outside our awareness, then the cueing frequency effect should appear in all participants, even in those who did not feel that they succeeded in following the instructions, at least for the visual task. On the contrary, if attentional refreshing is a more strategic and controlled process, and because attentional refreshing is domain-general, its effect should have been present in participants who reported to have actively followed the instructions in the verbal domain. Our data are not congruent with any of these accounts. We had evidence against an effect of cueing manipulation with participants that subjectively felt that they followed that task in the verbal part of Experiment 2, and against an effect of cueing manipulation on participants that felt that they did not succeed in following the task in the visual part of Experiment 2.

Alternatively, responses about following the instructions in the visual task can be based on working memory content, and only participants who had access to memory traces of colors in working memory thought they followed instructions. This supports the proposal that working memory content is accessible to consciousness (Baars & Franklin, 2003; Baddeley, 2000), and is in line with Rademakers, Tredway and Tong (2012) who have shown through another paradigm that young adults are able to introspect whether they had a memory representation for what was previously presented at a cued location. Thus, compliance to the instructions in the visual task could have been mistaken by participants with memory accessibility. However, this does not explain why subjective assessment of instructions compliance was not related to cueing frequency benefit and memory accessibility in the verbal domain. Once more, this supports the functional differences between the two domains.

Another possible explanation for the divergence of the cueing frequency benefit between the visual and verbal domains could be linked to differences in strategies implemented by the participants. Indeed, in Experiment 2, the majority of participants reported using some kind of elaboration strategy to maintain the words in the verbal

task, while a similar strategy was never reported for maintaining colors in the visual task. Elaboration is known to improve recall performance at short term (Campoy & Baddeley, 2008; Engle, Nagle, & Dick, 1980; Haarmann & Usher, 2001; Nishiyama, 2018; Shivde & Anderson, 2011, but see Bartsch et al., 2018), as it could create episodic cues that helps retrieval from long-term memory. The putative effect of attentional refreshing for verbal material could have been masked by the use of an elaboration strategy and the reliance on long-term memory. However, we think this is rather unlikely. Indeed, we implemented in Experiment 3 a manipulation that varied the reliance on long-term memory by presenting either words or non-words. The latter would benefit less from information stored in long-term memory, and the use of elaboration is less likely to provide efficient cues. We nevertheless found evidence for an absence of the effect in both types of memory items.

Finally, our preferred account is that the cues have a different impact on verbal and visual items because their memory traces are not stored in the same memory system and retrieval processes are of different nature. The cueing frequency benefit in the visual task might reflect the enhancement of visual traces stored primarily in an iconic memory, while verbal items are stored in working memory as distinct memory traces temporally organized. It might be assumed that when colors are presented on screen, this stimulus is stored as a mental image in sensory memory (e.g., Atkinson & Shiffrin, 1968). The effect of the arrows is to direct attention to some features of this mental image (Cowan, 2001), with the attended elements being encoded in working memory, facilitating their maintenance. Hence, at recall, cued elements are better recalled than non-cued, because they benefit from working memory encoding. The small size of the cueing frequency benefit also suggests that this attentional focusing succeeds only sporadically. Increasing the number of cues towards the same spatial region increases the probability of success of the transfer in working memory, resulting in a slightly better performance. Indeed, it should not be forgotten that in Experiment 2, recall error in color reconstruction task progressed from 71° with no cue to 69° and 64° with one and two cues, respectively, which is a very small

improvement for a large error. For a concrete example, when presenting a yellow, participants would choose, on average, a green or a red (the error can be in both directions in the color wheel). Hence, the manipulation of the cueing frequency would impact recall without necessarily involving attentional refreshing.

By contrast, in the verbal task, each letter, word or non-word would be encoded into working memory as a unique and distinct representation, most probably of phonological nature, this encoding being facilitated by their sequential and slow presentation. Their maintenance in working memory can be done through either domain-specific verbal rehearsal or attentional refreshing, though our results point toward a maintenance through verbal rehearsal. Both mechanisms are sequential in nature, items being maintained in sequential order (Camos, Lagner & Loaiza, 2018). At recall, the sequential order of memory traces is mapped onto the spatial arrangement of the boxes to retrieve the target item by scrolling the memorized list. This is facilitated by the fact that the items were always presented on screen in the same clockwise order starting by the top box. In other words, verbal items are maintained in a sequential order, and spatial cueing cannot facilitate their maintenance because the schedule of arrows appearance does not fit, and even might interfere with, the sequential progression of the working memory maintenance induced by the paradigm. This may explain the trend for worse performance in cued conditions compared to baseline. This could also explain why articulatory suppression had such a large effect on performance (a drop of 40% in Exp. 1) because it disrupts verbal rehearsal known to its critical role in order maintenance (e.g., Baddeley, 1986; Tan & Ward, 2008).

6.9 Conclusion

Contrary to what is observed in the visual domain, we found consistent evidence against an effect of the number of retro-cues in the verbal domain. Scrutiny on previous findings supports this divergence between the two domains. Together, this goes against the idea that attentional refreshing is implicated in the cueing benefit

reported in the guided refreshing paradigm, and that this paradigm is a direct test of attentional refreshing.

Chapter 7: General discussion

The aim of this thesis was to investigate the influence of semantic long-term memory on attention-based maintenance processes in working memory. In the introduction, we saw that two attention-based processes are involved in the maintenance of information in working memory, namely consolidation and attentional refreshing. Consolidation is the process that transforms fleeting iconic traces into more stable working memory representation (see Ricker et al. 2018 for a review), whereas attentional refreshing is used to maintain information in working memory by focusing attentional resources on them (see Camos et al. 2018 for a review). In experimental chapter 4.1 and 4.2, we manipulated the familiarity of visuospatial memoranda orthogonally to attentional refreshing opportunity. In chapter 5, consolidation opportunity was manipulated orthogonally to the familiarity of visuospatial memoranda. Finally, in experimental chapter 6, we tried to replicate results from a novel paradigm, called the guided refreshing paradigm (Souza, Rerko & Oberauer, 2015), which is thought to directly manipulate controlled attentional refreshing. Our results are examined in the light of our general hypothesis. Possible limits of our experimental paradigm and possible new paths for future research are discussed.

7.1 The Impact of Semantic Long-Term Memory on Attention-Based Processes in Working Memory

Each experimental chapter had at least one experiment where semantic long-term memory was manipulated in addition to swift attentional refreshing opportunity (chapter 4.1 and 4.2), consolidation (chapter 5), and controlled attentional refreshing (chapter 6). The results of these experiments are discussed with a focus on the impact of long-term memory factors in human tasks.

7.1.1 Semantic Long-Term Memory and Swift Attentional Refreshing functioning

The literature on swift attentional refreshing showed that not every type of memoranda was refreshable. Although verbal material has always been shown to be susceptible to attentional refreshing opportunity manipulation (Barrouillet et al., 2007; Vergauwe, Barrouillet & Camos, 2010; Vergauwe, Dewaele, Langerock & Barrouillet, 2012), it seems that some specific visuospatial material are immune to the effect of attentional refreshing manipulation (Ricker & Cowan, 2010; Ricker & Vergauwe, 2020; Vergauwe, Camos & Barrouillet, 2014). This shows that attentional refreshing is not as ubiquitous as previously thought. It also raises the question as to why some type of memoranda cannot be refreshed.

We tested the hypothesis that attentional refreshing relies on long-term memory representation in a process similar to redintegration: attentional refreshing would use already existing information in long-term memory to recreate what has been forgotten from the information held in working memory due to temporal decay and interference. If this hypothesis was true, then more familiar memoranda should benefit more from attentional refreshing opportunity than less familiar memoranda. This would have been marked by a statistical interaction, where the difference between high and low familiarity memoranda would be greater in conditions with more opportunity for refreshing (i.e., low cognitive load conditions), compared to condition with less opportunity to refresh (i.e., high cognitive load conditions). In low cognitive load condition, high familiarity memoranda would benefit more from the opportunity to refresh, and would thus be refresh to a higher degree than less familiar memoranda.

In the verbal domain, a recent study showed that, although both attentional refreshing opportunity and memoranda familiarity influence memory performance, both effects are separated (Camos et al., 2019). These authors concluded that, at least in the verbal domain, attentional refreshing did not rely on semantic long-term

memory. Although not congruent with our hypothesis, this study did not completely rule it out. Memoranda familiarity was manipulated via the lexical frequency and lexicality effect (higher lexical frequency words are better recalled than lower lexical frequency words, and real words are better recalled than pseudo-words, respectively). These effects rely on the fact that more frequent words (or real words) have better long-term memory representation than less frequent words (or pseudo-words), because they are encountered more frequently in our day-to-day lives. But, having a differential in representational stability in long-term memory does not mean that no representation at all exists in long-term memory. Although less familiar, lower lexical frequency words are still encountered in daily lives and thus have a representation in long-term memory. Regarding the lexicality effect, the pseudo-words used in the study from Camos and colleagues (2019) still followed the phonotactical frequency of the source language (here: French)⁵. Since the phonotactical frequency of a language is also stored in long-term memory (Bonte et al. 2005; Gonzalez-Gomez, Poltrock & Nazzi, 2013), it may be that attentional refreshing used this information in the same way for words and pseudo-words, and thus no difference was detected between both types of words in regard of attentional refreshing manipulation. Due to this possibility, we aimed at testing the hypothesis of an involvement of semantic long-term memory on attentional refreshing in the visuospatial domain.

In chapter 4.1, we manipulated the familiarity of visuospatial memoranda orthogonally to attentional refreshing opportunity. We found a similar pattern of results to the Camos and colleague (2019) study. Memoranda familiarity as well as cognitive load manipulation had an impact on memory task performance in the predicted direction. More familiar memoranda induced better memory task performance and higher cognitive load induced worst memory task performance.

⁵ The phonotactical frequency of a language corresponds to the probability of a phoneme to be followed by another phoneme in a specific language. For example, "ba" have a higher phonotactical frequency in French than "po", because "ba" can be found in more words than "po".

However, we found evidence against an interaction between both factors. Memoranda were refreshed at the same degree irrespective of their intrinsic familiarity. This contradicts the possibility that the effect between both factors was masked in the verbal domain and argues against our general hypothesis. One last possible confound could have been that the absence of the hypothesized effect was due to the specific memoranda we used in this study.

In chapter 4.2, we thus tried to replicate our pattern of results, but this time we used faces as memoranda, instead of drawings. To manipulate long-term memory, we relied on the Own-Race effect (faces from our own ethnicity induce better recall performance than face from another ethnicity), which have been shown to rely on representational difference in long-term memory. Although we found an effect of face ethnicity in the expected direction, the results showed that faces were not even influenced by attentional refreshing opportunity manipulation. This led to the hypothesis that faces cannot be refreshed. Since our research question relied on the conjoint impact of attentional refreshing and familiarity in working memory, we did not follow up on chapter 4.2. Nonetheless, the fact that the own-race effect was present in our data, but there was evidence against an effect of attentional refreshing opportunity manipulation also argue against our general hypothesis. It shows that the impact of long-term memory factor on working memory performance is independent of the presence or not of an effect of attentional refreshing opportunity.

Taken together, results from Camos and colleagues (2019) and our results from chapter 4.1 and 4.2 argue against the hypothesis that more familiar memoranda are refreshed more efficiently than less familiar memoranda, irrespective of the memoranda domain. We gathered congruent evidence that, despite both effects being present in our data, visuospatial familiarity and swift attentional refreshing opportunity affect memory task performance independently from one another. This means that the locus of familiarity effects in working memory is not situated during maintenance. This is congruent with the proposition from Vergauwe and Cowan

(2015) which suggest that attentional refreshing works via the rapid scanning of the content of working memory (what is currently in the focus of attention), and against the hypothesis that attentional refreshing uses a redintegration-like process.

7.1.2 Guided refreshing paradigm, controlled attentional refreshing and semantic long-term memory

Regarding controlled refreshing, our results from chapter 6 contradict previous findings from Souza and colleagues (2018) in the verbal domain. The guided refreshing paradigm was initially developed as a way to directly manipulate attentional refreshing in working memory. In this paradigm, several cues appeared in between item presentation and recall, and participants were tasked to “think of” the memoranda presented where the retro-cues pointed (Souza et al. 2014). Since controlled refreshing is defined as the act of consciously thinking of a memoranda, the authors thus argued that this paradigm directly manipulated controlled attentional refreshing. This paradigm is an evolution of the simple retro-cue paradigm.

Former studies showed that a single retro-cue appearing during maintenance impacted memory task performance: the memoranda which have been retro-cued are better recalled than the ones that have not been retro-cued (Griffin & Nobre, 2003; Landman et al. 2003, see Souza & Oberauer, 2016 for a review). The guided refreshing expands on this paradigm by including multiple retro-cues during maintenance, instead of only one. In the visuospatial domain, Souza and colleagues (2014, 2018) found a monotonic relation between the number of cueing and memory task performance in a continuous color reconstruction task. The more a memorandum was cued, the better the memory performance for this memorandum was.

In experimental chapter 6, we effectively replicated this effect in the visuospatial domain (Experiment 2). However, we found contradictory results in the verbal

domain. Souza and colleagues (2018) found evidence for an effect of the number of retro-cues on memory task performance in the verbal domain, although much more ambiguous than for the visuospatial domain. In contrast, we found, at best, ambiguous evidence against an effect of the number of cueing on memory performance in the verbal domain. Reanalysis of the data from Souza and colleagues (2018) also showed that the effect of retro-cues on memory performance in the verbal domain was mostly driven by a difference between conditions where memoranda were not cued compared to when they were cued only once. There was inconclusive evidence that two retro-cues enhanced recall performance compared to a condition where the memoranda were cued only once.

This casted doubt on the implication of attentional refreshing in this paradigm. Indeed, their argument for an implication of controlled attentional refreshing in their paradigm was that each cue would enhance memory performance. If the effect is only driven by a difference between cued and not-cued memoranda, then the guided refreshing paradigm is not different than a simple retro-cue paradigm. The single retro-cue effect can be straight forwardly explained by a prioritization of some information compared to other in working memory. It has been shown that specific memorandum can be prioritized in working memory (Atkinson et al. 2021; Kumar et al., 2013; Myers, Stokes & Nobre, 2017), usually at the cost of memoranda that are not cued. Since attentional refreshing is thought to work serially on all memoranda, the fact that prioritization impacts not-cued items goes against the idea that this benefit stems from better refreshing to the cued items compared to the not-cued memoranda. If it was the case, not-cued memoranda should attain the same recall performance than in baseline trials, where the maintenance interval stayed unfilled. Results showed that performance to the memory task are worse for memoranda that are not being cued, compared to baseline trials.

Furthermore, we also manipulated semantic long-term memory in the guided refreshing with verbal material (chapter 6, experiment 3). We manipulated words

lexicality and the number of cueing, and only found an effect of lexicality on recall performance. We found evidence against an effect of cueing manipulation on recall performance. This pattern of results also argues against an impact of long-term memory on controlled refreshing because the effect of lexicality was present even though retro-cues did not impact recall performance. Thus, even if we suppose that guided refreshing actually manipulates directly controlled attentional refreshing and the effect was masked in the verbal domain for whatever reason, the fact that lexicality effect was present shows that this effect emerged irrespective of an implication of controlled attentional refreshing.

To conclude, our results are congruent with the result of chapter 4.1 and 4.2, and show that controlled, as well as swift, attentional refreshing are not impacted by semantic long-term memory factors. This means that the locus of the familiarity effect in working memory has to be found elsewhere.

7.1.3 Semantic Long-term Memory and Consolidation in Working Memory

Contrary to the literature on attentional refreshing, previous studies on consolidation in working memory already investigated whether this process was influenced by semantic long-term memory. Both studies that investigated this hypothesis showed that more time for consolidation enhanced memory task performance (Blalock, 2015; Xie & Zhang, 2018). The authors of these studies thus concluded that more familiar memoranda were consolidated more quickly than less familiar memoranda.

However, both studies manipulated consolidation as the delay between memoranda presentation and the appearance of a mask (or the start of a concurrent task, i.e., pre-mask delay manipulation). As we saw in the theoretical introduction of this thesis, manipulating consolidation using a pre-mask delay manipulation makes it difficult to disentangle whether the manipulation impacted encoding processes, consolidation, or both. In addition, the authors of both studies argue that familiarity

impacted the speed of consolidation, without showing evidence for a direct impact of consolidation delay on reaction time to the concurrent task. In chapter 5, we manipulated consolidation using a post-mask delay manipulation (consolidation operationalized as the delay between a mask directly following memoranda presentation and the start of a concurrent task) and measured more directly the speed at which consolidation was done, via measuring the postponement induced by consolidation on reaction time to a concurrent task. We found an effect of post-mask delay manipulation in the direction we hypothesized. Shorter delay induced slower reaction time to the concurrent task in Experiment 1, and also induced worse memory task performance in Experiment 2. Taken together, this pattern of results is congruent with the supposition that longer consolidation delay enhanced memory task performance, compared to shorter delay.

Although the impact on memory performance was not evident in Experiment 1, this does not impede the conclusion that consolidation was effectively manipulated by our post-mask delay manipulation. Indeed, past studies showed that consolidation delay manipulation only impact recall performance when the task difficulty is high overall (Nieuwenstein & Wyble, 2014). This is congruent with the supposition that memoranda need to attain a certain level of consolidation to be effectively recalled. It may be that this threshold also depends on the recall procedure. Since recognition can rely on more degraded representation than recall, memoranda can be consolidated to a lesser degree and still be correctly recognized, but the same level of consolidation would not be enough if recall procedure was used. This would explain why we detected an effect of consolidation delay on recall performance in Experiment 2, but not in Experiment 1. Experiment 2 used a recall procedure, whereas Experiment 1 used a recognition procedure.

We also found an effect of familiarity on reaction time to the concurrent task in Experiment 1, where more familiar memoranda induced shorter reaction time than less familiar ones. Since we showed that consolidation was manipulated via our post-

mask delay, the impact of familiarity on the reaction time to the concurrent task argues in favor of an effect of memoranda familiarity on consolidation, where more familiar memoranda are consolidated more rapidly than less familiar memoranda. This would mean that, for a constant consolidation delay, more familiar memoranda could make a better use of it than less familiar memoranda, and thus would be consolidated to a higher degree than less familiar memoranda. This difference in consolidation strength would then explain, at least in part, the memory advantage for more familiar memoranda. To our knowledge, Experiment 1 from chapter 5 is the first direct evidence that more familiar memoranda are effectively consolidated more quickly than less familiar memoranda, thanks to our post-mask delay manipulation and our measure of reaction time to the processing task. Nonetheless, although we found the expected impact of memoranda familiarity on consolidation, familiarity likely affect other processes in working memory. This is discussed in the next section.

7.1.4 The Locus of Long-term familiarity Effect in Working Memory

In the experiments presented in this thesis that directly manipulated familiarity, we almost always found an effect of memoranda familiarity on memory task performance. It was only in chapter 5, Experiment 2 that our familiarity manipulation did not lead to an effect on recall performance. As indicated above in our discussion of attentional refreshing and semantic long-term memory (chapter 7.1.1), we ruled out the possibility that familiarity effects in working memory stem from attentional refreshing functioning. This is not the case for consolidation. We gathered evidence that consolidation is influenced by semantic long-term memory factors. However, it is hard to conclude that the whole effect of memoranda familiarity on recall performance stems from the consolidation process alone.

In chapter 5, experiment 1, we found that more familiar memoranda were consolidated more quickly than less familiar memoranda. However, the mean difference between both kind of memoranda was around 22ms, which correspond to a mean increase in reaction time of around 3.6% induced by our familiarity

manipulation on the concurrent task's reaction time. Comparatively, the impact of familiarity on recall performance was bigger. More familiar memoranda induced an increase of 11% compared to the recall score of less familiar memoranda.

Thus, it seems rather unlikely that the impact of familiarity on consolidation explains the whole familiarity advantage on recall performance. Since it has been shown that presentation duration impacts the size of familiarity effect in working memory (Malmberg & Nelson, 2003; Zhou, Mondloch & Emerich, 2018), familiarity likely also impacts encoding processes in working memory, with more familiar memoranda being encoded more quickly than less familiar one. Evidence for the impact of familiarity on encoding can be gathered if we compare our experiments in chapter 6 with studies from Blalock (2015) and Xie and Zhang (2018). In both these studies, the authors found an interaction between familiarity and their pre-mask delay manipulation. In contrast, we found evidence against the interaction between familiarity and post-mask delay manipulation. Since pre-mask delay manipulation manipulated encoding and consolidation concurrently, but post-mask delay manipulation only manipulate consolidation, the overall pattern of result is congruent with the supposition that familiarity impact encoding processes, with more familiar memoranda being encoded more efficiently than less familiar memoranda. This difference in encoding efficiency would lead to a difference on the strength of encoding, and thus on recall performance.

In addition to an effect of memoranda familiarity during encoding and consolidation, familiarity could also impact memory performance at recall. This impact at recall could be two-fold. Firstly, since we mostly used serial recognition, participants could also partly rely on a feeling of familiarity to sort presented memoranda and non-target (Gardiner & Parkin, 1990). It has been shown that this "feeling of familiarity" relies on a neuronal network distinct from direct recollection (Duarte, Ranganath, Winward, Hayward & Knight, 2004), and that it can be used to enhance memory performance. Thus, familiarity also impacts memory task

performance at retrieval: more familiar memoranda could induce a stronger feeling of familiarity, compared to less familiar memoranda. More familiar memoranda would then be more easily recognized from non-targets compared to less familiar memoranda, and this would enhance memory task performance for the higher familiarity memoranda.

Secondly, memoranda familiarity could also impact memory task performance at retrieval via a redintegration process (Hulme et al., 1997). More familiar memoranda could be reconstructed more easily at retrieval than less familiar memoranda, because more familiar memoranda have a better long-term representation to reconstruct the trace compared to less familiar memoranda.

It thus seems that familiarity effect in working memory arises as a conjunction of an impact of long-term memory factor on encoding, consolidation and at retrieval. More familiar memoranda are encoded and consolidated faster than less familiar memoranda, and thus can make a better use of the encoding and consolidation delays. In addition, more familiar memoranda would induce a bigger feeling of familiarity than less familiar memoranda, which would also enhance memory task performance at retrieval for higher familiarity memoranda compared to lower familiarity memoranda.

Semantic long-term memory thus impacts the functioning of working memory at several level. As we saw above, it seems that long-term memory impact the creation of working memory representation (or the entrance into the focus of attention, depending on the theoretical background), but not their maintenance via attentional refreshing. Although one of the goal of this thesis was to add new data that could inform the debate on the separation of working memory and long-term memory, the overall results are not clear on this regard.

If we consider that working memory and long-term memory are separated (as in the TBRS model, chapter 1.2.3, and the Multicomponent model of working memory, chapter 1.2.1, then our results can be explained by the fact that long-term memory

impact the creation of the working memory representation, but then, once in working memory, long-term memory does not impact the representation anymore. Only when the information has to be given back, then long-term memory influence memory performance again.

In contrast, if we consider that working memory is the activated part of long-term memory (as in the embedded-processes model of working memory, chapter 1.2.2), then long-term memory impact the entrance of information into the focus of attention (i.e., impact encoding and consolidation), but not the actual maintenance process. This would mean that attentional refreshing works by reactivating the traces in the focus of attention, irrespective of how the trace is actually represented in the focus of attention.

Overall, the link between long-term memory and working memory goes further than a simple impact at a specific level. This shows a high level of interaction between both constructs. Future work should aim at more precisely investigate the degree at which long-term memory impact each of the working memory steps (encoding, consolidation and retrieval). This would inform more precisely the actual relation between both constructs.

7.2 Limits, future studies and methodological advancement

In this thesis, we used three main experimental tasks: complex-span tasks (chapter 4.1, 4.2 and 5), Brown-Peterson (chapter 4.2), and guided refreshing (chapter 6). Complex-span tasks have been used extensively to investigate attention-based maintenance process in working memory (Barrouillet et al. 2004, 2007; Camos et al. 2019; Vergauwe et al. 2010). Indeed, the impact of a cognitive load manipulation on recall performance has been widely replicated. This cognitive load effect has been used to evidence the implication of attentional refreshing for the maintenance of information in working memory. However, it is important to note that complex-spans do not manipulate refreshing directly but manipulate the opportunity to engage in attentional refreshing. The same criticism can be raised about the Brown-Peterson paradigm. In the latter, the opportunity to engage in attentional refreshing is manipulated via the number of memoranda to refresh, in contrast to the complex-span which manipulates the duration during which participants can engage in attentional refreshing.

Nonetheless, we showed that complex-span could also be useful to investigate the memory load effect on concurrent task reaction times. In chapter 4, in Experiments 2A and 2B, we found the expected effect of memory load on the reaction time to the concurrent task with the use of a complex-span paradigm instead of the Brown-Peterson task developed by Vergauwe and colleagues (2014). Indeed, we found the overall same postponement of 30-50ms induced by each new item held in working memory on reaction times to a concurrent task. This effect of memory load on reaction time in a complex-span was also evident in both experiments of chapter 5, although to a lesser degree. This could be explained because, due to the overall design of the tasks, it was more difficult to parse out trials in which participant did not completely succeed on the memory task from trials in which they were successful. In not completely successful trials, it may be that some memoranda were not refreshed at all and thus did not induce a postponement on the

reaction time to the concurrent task. Hence, reaction time in these trials should have been shorter than in trials in which all memoranda were refreshed, which, in turn, would have artificially lowered our estimate of refreshing speed.

Regarding paradigm we used in chapter 6, the guided refreshing paradigm was developed in the aim of directly manipulating attentional refreshing. However, as we saw in preceding sections of this thesis, it is not clear whether the guided refreshing paradigm does manipulate attentional refreshing. Our results, although more ambiguous, argue against this possibility. The aim of a direct path to manipulate attentional refreshing is thus still underway

7.2.1 Limits to our visuospatial familiarity manipulation

One possible limitation to our results is linked to how we operationalized visuospatial familiarity. Contrary to the verbal domain where words familiarity can be indexed by the lexical frequency or the lexicality of the words, there is no such index already existing in the visuospatial domain. In experiment 4.1 and 5 (Experiment 1), familiarity was manipulated by creating two pools of memoranda that we thought differed in familiarity. The first group consisted of drawing from real-life objects taken from the Snodgrass and Vanderwart (1980) image set. The un-familiar image set was created using smoothly connected features from different images from the Snodgrass and Vanderwart image set (Soldan et al. 2008). This created images that represented nothing specific but was as visually complex as the drawing of real objects used in the familiar image set.

In our experiments, drawings from the “familiar” group always induced better memory task performance than drawings taken from the un-familiar group. Although we interpreted this difference as an impact of memoranda familiarity on recall performance, we cannot completely rule out the implication of possible labelling effect in our results. Indeed, more familiar memoranda were also easier to name than less familiar memoranda. Since labelling enhances memory task performance in

working memory (Souza, Overkott & Matyja, 2021), it may be that the familiarity effect we detected was due to a labelling effect instead.

Nonetheless, we argue that the way we manipulated visuospatial familiarity did in fact manipulate familiarity, and for two reasons. First, participants were always under articulatory suppression in our experiments. This limited their opportunities to engage in verbal maintenance, like verbal rehearsal, and thus limited the possibility to maintain only the name of the drawings. Secondly, the name we gave to an object is part of its overall long-term memory representation. A difference in long-term memory can also imply a difference in the capacity to name the object. Thus, even if verbal labelling did induce some part of the measured effect, it does not mean that long-term memory was not effectively manipulated with our image manipulation.

In chapter 5, Experiment 2, we manipulated familiarity more directly, by training a subset of participants on the actual figures used as memoranda. Our result showed no effect of this manipulation, neither on reaction time to the concurrent task nor on recall performance in the complex-span. But, as we saw in the section dedicated to the analysis of this experiment (chapter 5.7), the absence of familiarity effect can be explained by the actual methodology we used. In future studies. Instead of manipulating familiarity between-subject, it might be interesting to manipulate familiarity within-subject by training participants only on a subset of figures, as in Blalock (2015). The performance between trained and un-trained figures would then be compared. This would allow to control inter-individual differences between participant samples. In addition, since the effect of memoranda familiarity on reaction time in Experiment 1 of chapter 5 was rather small, using a within-subject manipulation would maybe allow to detect smaller difference than a between-subject experimental plan.

7.3 Conclusion

In a series of four experimental studies, we investigated whether attention-based process during working memory maintenance were influenced by semantic long-term memory factors. Regarding attentional refreshing, our results in the visuospatial domain were similar to those in the verbal domain and showed a clear separation between attentional refreshing and semantic long-term memory. Memoranda which had better long-term memory representation were not refreshed more efficiently than memoranda that had worse long-term memory representation, irrespective of the memoranda domain.

In contrast, we found that consolidation in working memory was influenced by semantic long-term memory factors. More familiar memoranda were consolidated more quickly than less familiar memoranda. However, this difference in consolidation speed did not always induce an effect on recall performance. The effect seemed to depend, at least in part, on the recall procedure used. When using recognition procedure, memoranda were consolidated up to a certain threshold which was sufficient to be effectively recognized, and that both familiar and un-familiar memoranda were consolidated above this threshold. In contrast, consolidation manipulation had an impact on the memory task performance when participant had to recall the memoranda. This is congruent with the hypothesis that consolidation works on a more continuous manner, instead of an "all-or-nothing" process.

Finally, we found evidence against the implication of attentional refreshing in the guided refreshing paradigm. Although first studies using this paradigm were promising (Souza et al. 2014, 2018), the fact that we found, at best, ambiguous effect of multiple retro-cues on recall performance argue against an implication of attentional refreshing on this effect.

Taken together, our results argue in favor of an effect of semantic long-term memory on encoding, consolidation and at retrieval, but not on attentional refreshing. Thus, the familiarity effect in working memory seems dependent on

several different processing steps. This shows a pervasive link between long-term memory and working memory, which goes further than a simple impact on one single working memory process. Future studies should aim at finding the limit, for each of this process, at which the familiarity effect emerges or disappears. This could inform more precisely the link between semantic long-term memory and working memory.

Bibliography

Aggleton, J. P. (2008). Understanding anterograde amnesia: Disconnections and hidden lesions. *The Quarterly Journal of Experimental Psychology*, *61* (10), 1441–1471. doi:10.1080/17470210802215335

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning & Verbal Behavior*, *22* (3), 261–295. doi: 10.1016/S0022-5371(83)90201-3

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). *An integrated theory of the mind*. *Psychological Review* *111*, (4). 1036-1060.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, *30*, 221–256

Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. (2021). Can valuable information be prioritized in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47* (5), 747–764. doi: 10.1037/xlm0000979

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1977). Human memory: a proposed system and its control processes. *Human Memory*, 7–113. doi: 10.1016/b978-0-12-121050-2.50006-5

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of experimental psychology. Human perception and performance*, *24* (3), 780–790. doi: 10.1037//0096-1523.24.3.780

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*, 166–172.

Baddeley, A. (1986). *Working memory*. Oxford, UK: Clarendon Press/Oxford University Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4 (11), 417–423. doi:10.1016/s1364-6613(00)01538-2

Baddeley, A. D. (1993). Working memory or working attention? In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 152–170). Clarendon Press/Oxford University Press.

Baddeley, A. D., & Lewis, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society*, 22, 403-405.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (p. 28–61). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781139174909.005

Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, 49 (6), 1393–1400. doi: 10.1016/j.neuropsychologia.2010.12.042

Baddeley, A. D., Hitch, G., & Allen, R. (2021). A multicomponent model of working memory. In: *Working memory: the state of the art*. Logie, R.H., Camos, V., & Cowan, N. (eds.). Oxford (UK): Oxford University Press, 10-43

Baddeley, A.D., & Hitch, G. (1974). Working Memory. In: Bower, G.H., Ed., *The Psychology of Learning and Motivation: Advances in Research and Theory* (p.47-89). New-York, Academic Press. doi: 10.1016/S0079-7421(08)60452-1

Bahrick, H. P., Bahrick, P. O., & Wittinger, R. P. (1975). Fifty years of memory for names and faces: a cross-sectional approach. *Journal of Experimental Psychology: General*, 104, 54-75.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. Doi: 10.1037/a0033872

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133* (1), 83–100. doi: 10.1037/0096-3445.133.1.83

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of experimental psychology. Learning, memory, and cognition*, *33* (3), 570–585. Doi: 10.1037/0278-7393.33.3.570

Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language*, *45* (1), 1–20. doi: 10.1006/jmla.2001.2767

Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*, *21*, 413–419.

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Hove, UK: Psychology Press.

Barrouillet, P., & Camos, V. (2021). The Time-Based Resource-Sharing Model of Working Memory in: *Working Memory: the state of the art*. Logie, R.H., Camos, V., & Cowan, N. (eds.). Oxford (UK): Oxford University Press, 85-115.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological review*, *118* (2), 175–192. doi: 10.1037/a0022324

Bartsch, L. M., & Oberauer, K. (in press). Reactivated visual masks do not disrupt serial recall. A failed Replication of Rey et al. 2018. *Experimental Psychology*. doi: 10.31234/osf.io/ftme7

Bartsch, L. M., Singmann, H & Oberauer, K. (2018). The effects of refreshing and elaboration on working memory performance, and their contributions to long-term memory formation. *Memory & Cognition*, *46*, 769-808.

Bayliss, D. M., Bogdanovs, J., & Jarrold, C. (2015). Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *Journal of Memory and Language*, *81*, 34-50

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

Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, *11* (10), 6–6. doi:10.1167/11.10.6

Belletier, C., Doherty, J., Jaroslawska, A., Rhodes, S., Cowan, N., Naveh-Benjamin, M., Barrouillet, P., Camos, V., & Logie, R. (in revision). Strategic adaptation to dual-task in verbal working memory: Potential routes for theory integration.

Blalock L. D. (2015). Stimulus familiarity improves consolidation of visual working memory representations. *Attention, perception & psychophysics*, *77* (4), 1143–1158. doi: 10.3758/s13414-014-0823-z

Blalock, L. D. (2013). Mask similarity impacts short-term consolidation in visual working memory. *Psychonomic Bulletin & Review* *20*, 1290–1295.

Blalock, L.D., & Clegg, B.A. (2010). Encoding and representation of simultaneous and sequential arrays in visuospatial working memory. *Quarterly Journal of Experimental Psychology*, *63* (5), 856-862, doi: 10.1080/17470211003690680

Bonin, P., Méot, A., Ferrand, L., & Roux, S. (2011). L'imageabilité: Normes et relations avec d'autres variables psycholinguistiques. *L'Année Psychologique*, *111*, 327-357.

- Bonte, M. L., Mitterer, H., Zellagui, N., Poelmans, H., and Blomert, L. (2005). Auditory cortical tuning to statistical regularities in phonology. *Clinical Neurophysiology*, *116*, 2765–2774. doi: 10.1016/j.clinph.2005.08.012
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105* (38), 14325-14329. doi: 10.1073/pnas.0803390105
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.
- Broadbent, D. E., & Broadbent, M. H. (1987). "From detection to identification: response to multiple targets in rapid serial visual presentation". *Perception and Psychophysics*. *42* (2), 105–13. doi:10.3758/bf03210498
- Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*. *10* (1), 12–21. doi:10.1080/17470215808416249.
- Brown, R. (1958). *Words and things*. New York, NY: Free Press.
- Camos, V. (2017). Domain-specific versus domain-general maintenance in working memory: Reconciliation within the time-based resource sharing model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 135–171). Elsevier Academic Press.
- Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A. & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New-York Academy of Science*, *1424* (1), 19-32. doi:10.1111/nyas.13616.
- Camos, V., Lagner, P. & Loaiza, V. (2017). Maintenance of item and order information in verbal working memory. *Memory*, *8*, 953-968.
- Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457-469.

Camos, V., Mora, G., & Oberauer, K. (2011). Adaptive Choice Between Articulatory Rehearsal and Attentional Refreshing in Verbal Working Memory. *Memory and Cognition*, *39*, 231-244.

Camos, V., Mora, G., Oftinger, A.-L., Mariz-Elsig, S., Schneider, P. & Vergauwe, E. (2019) Does semantic long-term memory impact refreshing in verbal working memory? *Journal of experimental psychology: Learning, Memory and Cognition*. *45* (9), 1664-1682. doi: 10.1037/xlm0000657

Camos, V., Mora, G., Oftinger, A.-L., Mariz-Elsig, S., Schneider, P., & Vergauwe, E. (2019). Does Semantic Long-Term Memory Impact Refreshing in Verbal Working Memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *45*, 1664-1682.

Campoy, G., & Baddeley, A. (2008). Phonological and semantic strategies in immediate serial recall. *Memory*, *16*, 329–340.

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*. *8*(4), 293–332. doi:10.1207/s1532690xci0804_2.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, *81*(1), 1–8. doi: 10.1016/j.jebo.2011.08.009

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4* (1), 55–81. doi: 10.1016/0010-0285(73)90004-2

Chen, H., & Wyble, B. (2015). Amnesia for Object Attributes. *Psychological Science*, *26* (2), 203–210. doi:10.1177/0956797614560648

Chen, H., & Wyble, B. (2016). Attribute amnesia reflects a lack of memory consolidation for attended information. *Journal of Experimental Psychology: Human Perception and Performance*, *42* (2), 225–234. doi:10.1037/xhp0000133

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual review of psychology*, *62*, 73–101. doi: 10.1146/annurev.psych.093008.100427

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3* (1), 37–60. doi: 10.3758/bf03210740

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity by using an indirect face-matching measure. *Perception*, *31* (8), 985–994. doi: 10.1068/p3335

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *4*, 354–373

Conway, A. R. A., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183. doi:10.1016/S0160-2896(01)00096-4

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (2007). Variation in working memory: An introduction. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), *Variations in working memory* (pp. 3-17). Oxford, UK: Oxford University Press.

Conway, A.R.A., Kane, M. J., Bunting, M. F., Hambrick D. Z., Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12* (5), 769-786. doi: 10.3758/BF03196772

Cowan N. (2008). What are the differences between long-term, short-term, and working memory? In: W.S. Sossin, J.-C. Lacaille, V.F. Castellucci & S. Belleville (Eds.) *Progress in Brain Research*, *169*, 323–338. doi: 10.1016/S0079-6123(07)00020-9

Cowan N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, 24 (4), 1158–1170. doi: 10.3758/s13423-016-1191-6

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, 104 (2), 163–191. doi: 10.1037/0033-2909.104.2.163

Cowan, N. (1992). Verbal memory span and the timing of spoken recall. *Journal of Memory and Language*, 31, 668-684.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford (UK): Oxford University Press.

Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. doi: 10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.

Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current direction in psychological science*, 19 (1), 51-57

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, 51, 42-100.

Cowan, N., Morey, C. C., & Naveh-Benjamin, M. (2021). An embedded-processes approach to Working Memory in: Logie, R. H., Camos, V. & Cowan, N. (eds). *Working Memory*. Oxford (UK): Oxford University Press.

Cowan, N., Rouders, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: What does it take to make them work? *Psychological Review*, 119 (3), 480–499. doi: 10.1037/a0027791

Cowan, N., Sauls, J. S., & Elliott, E. M. (2002). The search for what is fundamental in the development of working memory. *Advances in Child Development and Behavior, 29*, 1-4.

Crist, R. E., Wu, C.-T., Karp, C., & Woldorff, M. G. (2008). Face processing is gated by visual spatial attention. *Frontiers in Human Neuroscience, 1* (10). doi: 10.3389/neuro.09.010.2007

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of experimental psychology. Human perception and performance, 35* (1), 94–107. doi: 10.1037/0096-1523.35.1.94

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466

Darling, S., Sala, S. D., & Logie, R. H. (2009). Dissociation between appearance and location within visuo-spatial working memory. *Quarterly journal of experimental psychology, 62* (3), 417–425. doi: 10.1080/17470210802321984

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychonomic bulletin & Review, 24*, 1651–1657. doi: 10.3758/s13423-017-1226-7

Den Heyer, K., & Barrett, B. (1971) Selective loss of visual and verbal information in STM by means of visual and verbal interpolated tasks. *Psychonomic Science 25*, 100–102. doi : 10.3758/BF03335861

Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Brain research. Cognitive brain research, 18* (3), 255–272. doi: 10.1016/j.cogbrainres.2003.10.010

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *Quarterly Journal of Experimental Psychology, 60*, 1227-1245.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review, 12*(6), 1127–1133.
doi:10.3758/bf03206454

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge University Press. doi: 10.1017/CBO9781139174909.007

Engle, R. W., Nagle, R. J., & Dick, M. (1980). Maintenance and generalization of a semantic rehearsal strategy in educable mentally retarded children. *Journal of Experimental Child Psychology, 30*, 438–454.

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? *Journal of Educational Psychology, 82* (4), 799–804. doi: 10.1037/0022-0663.82.4.799

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105* (3), 482–498. Doi: 10.1037/0033-295X.105.3.482

Freiwald, W., Duchaine, B., & Yovel, G. (2016). Face Processing Systems: From Neurons to Real-World Social Perception. *Annual review of neuroscience, 39*, 325–346. Doi: 10.1146/annurev-neuro-070815-013934

Frey, P. W., & Adesman, P. (1976). Recall memory for visually presented chess positions. *Memory & Cognition, 4* (5), 541-547

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition*, 18 (6), 579–583. Doi: 10.3758/BF03197100

Gardner, M. K., Rothkopf, E. Z., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory & Cognition*, 15 (1), 24–28. doi:10.3758/bf03197709

Gonzalez-Gomez, N., Poltrock, S., and Nazzi, T. (2013). A “bat” is easier to learn than a “tab”: effects of relative phonotactic frequency on infant word learning. *PLoS One* 8:e59601. doi: 10.1371/journal.pone.0059601

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15, 1176–1194.

Guitard, D., Saint-Aubin, J., & Cowan, N. (2021). Asymmetrical interference between item and order information in Short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 47 (2), 243-263. doi:10.1037/xlm0000956.

Haarmann, H., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review*, 8, 568–578.

Haberlandt, K. (1997). *Cognitive Psychology* (2nd edition). Boston, MA: Allyn & Bacon.

Hancock, P. J., Bruce, V., V, & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in cognitive sciences*, 4 (9), 330–337. doi: 10.1016/s1364-6613(00)01519-9

Higgins, J.A., & Johnson, M.K. (2009). The consequence of refreshing for access to non-selected items in young and older adults. *Memory & Cognition*, 37, 164-174.

Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2020). Attention and binding in visual working memory: Two forms of attention and two kinds of buffer storage. *Attention, perception & psychophysics*, 82 (1), 280–293. doi: 10.3758/s13414-019-01837-x

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30* (6), 685–701. doi: 10.1016/0749-596X(91)90032-F

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, M., & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: evidence for a redintegration process in immediate serial recall. *Journal of experimental psychology. Learning, memory, and cognition, 23* (5), 1217–1232. doi: 10.1037//0278-7393.23.5.1217

Jackson, M, C., & Raymond, J. E. (2008). Familiarity enhances visual working memory for faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 556-568.

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37* (3), 688–705.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26* (3), 683–702. doi: 10.1037//0278-7393.26.3.683

Johnson, A. J., & Miles, C. (2019). Visual Hebb repetition effects survive changes to both output order and concurrent articulation. *Journal of Cognitive Psychology, 31* (3), 276–284. doi: 10.1080/20445911.2019.1586715

Johnson, M. K. (1992). MEM: mechanisms of recollection. *Journal of Cognitive Neuroscience. 4*, 268–280.

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience, 5*, 339–361.

Johnston, R. A., & Edmonds, A. J. (2009) Familiar and unfamiliar face recognition: A review. *MEMORY, 17* (5), 577-596. doi: 10.1080/09658210902976969

Jolicoeur, P., & Dell'Acqua, R. (1998). The Demonstration of Short-Term Consolidation. *Cognitive Psychology, 36*, 138–202.

Jones, D. M., & Macken, B. (2018). In the Beginning Was the Deed: Verbal Short-Term Memory as Object-Oriented Action. *Current Directions in Psychological Science, 27*(5), 351–356. doi: doi.org/10.1177/0963721418765796

Joseph, T. N., & Morey, C. (In press). Impact of memory load on processing diminishes rapidly during retention in a complex span paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition*. Retrieved from <https://osf.io/dku3c/>

Kail, R. V. (2007). Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychological science, 18* (4), 312–313. doi: 10.1111/j.1467-9280.2007.01895.x

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130* (2), 169–183. doi: 10.1037/0096-3445.130.2.169

Kleinsorge, T., & Gajewski, P.D. (2004). Preparation for a forthcoming task is sufficient to produce subsequent shift costs. *Psychonomic Bulletin & Review 11*, 302–306. doi: 10.3758/BF03196574

Kowialieski, B., & Majerus, S. (2018). The non-strategic nature of linguistic long-term memory effects in verbal short-term memory. *Journal of Memory and Language*, *101*, 64–83

Krumm, S., Ziegler, M., & Buehner, M. (2008). Reasoning and working memory as predictors of school grades. *Learning and Individual Differences*, *18* (2), 248–257. doi: 10.1016/j.lindif.2007.08.002

Kumar, S., Joseph, S., Pearson, B., Teki, S., Fox, Z. V., Griffiths, T. D., & Husain, M. (2013). Resource allocation and prioritization in auditory working memory. *Cognitive neuroscience*, *4* (1), 12–20. doi: 10.1080/17588928.2012.716416

Kursawe, M. A., & Zimmer, H. D. (2015). Costs of storing colour and complex shape in visual working memory: Insights from pupil size and slow waves. *Acta Psychologica*, *158*, 67–77. doi: 10.1016/j.actpsy.2015.04.004

Landman, R., Spekreijse, H., & Lamme, V. A. (2003) Large capacity storage of integrated objects before change blindness. *Vision Research*; *43*, 149–164.

Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2018). What is the time course of working memory attentional refreshing? *Psychonomic bulletin & review*, *25* (1), 370–385. doi: 10.3758/s13423-017-1282-z

Lepsien, J., & Nobre, A. C. (2006). Cognitive control of attention in the human brain: Insights from orienting attention to mental representations. *Brain Research*, *11*, 20-31.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22* (6), 1364–1382. doi:10.1037/0278-7393.22.6.1364

Lewandowsky, S., & Farrell, S. (2000). A reintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychological Science*, *63*, 163-173.

Liang, F., & Xiong, M. (2013). Bayesian detection of causal rare variants under posterior consistency. *PLoS ONE*, *8*, e69633.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20* (4), 240–245. doi: 10.1177/0963721411415340

Logie, R., Camos, V., & Cowan, N. (2021). *Working Memory: The State of the Science*. Oxford (UK): Oxford University Press. doi: 10.1093/oso/9780198842286.001.0001

Logie, R.H., Camos, V., & Cowan, N. (2021). *Working Memory: State of the Science*. Oxford, UK : Oxford University Press.

Los, S. A., & Van der Burg, E. (2010). The origin of switch costs: task preparation or task application? *Quarterly journal of experimental psychology*, *63* (10), 1895–1915. doi: 10.1080/17470211003651849

Majerus, S., & Van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, *21* (2), 303–310. <https://doi.org/10.1348/026151003765264101>

Malmberg, K. J., & Nelson, T. O. (2003). The word frequency effect for recognition memory and the elevated-attention hypothesis. *Memory & cognition*, *31* (1), 35–43. doi: 10.3758/bf03196080

MATLAB. (2010). version 7.10.0 (R2017a). Natick, Massachusetts: The MathWorks Inc.

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, *58*, 480–494.

Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain and Cognition*, *52* (2), 250–257. doi:10.1016/S0278-2626(03)00081-2

Migo, E., Montaldi, D., Norman, K. A., Quamme, J., & Mayes, A. (2009). The contribution of familiarity to recognition memory is a function of test format when using similar foils. *Quarterly journal of experimental psychology*, *62* (6), 1198–1215. doi: 10.1080/17470210802391599

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. doi: 10.1037/h0043158

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118* (1), 43–71. doi:10.1037/0096-3445.118.1.43

Moore, C. D., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *26* (43), 11187–11196. doi: 10.1523/JNEUROSCI.1873-06.2006

Morey, C. C. (2018). The case against Specialized Visual-Spatial Short-Term Memory. *Psychological bulletin*, *144* (8), 849-883

Morey, C. C., & Cowan, N. (2018). Can we distinguish three maintenance processes in working memory? *Annals of the New York Academy of Sciences*, *1424* (1), 45–51. doi: 10.1111/nyas.13925

Morey, C. C., Cowan, N., Morey, R. D., & Rouder, J. N. (2011). Flexible attention allocation to visual and auditory working memory tasks: Manipulating reward induces a trade-off. *Attention, Perception, & Psychophysics*, *73*, 458–472.

Morey, R. D., & Rouder, J. N. (2018). Bayes Factor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2. <https://CRAN.R-project.org/package=BayesFactor>

Morton, J. (1979). Word recognition. In J. Morton & J. C. Marshall (Eds.), *Psycholinguistics volume 2—structures and processes* (pp. 107–156). London, UK: Paul Elek.

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends in cognitive sciences*, *21* (6), 449–461. doi: 10.1016/j.tics.2017.03.010

Myers, N. E., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2018). Benefits of flexible prioritization in working memory can arise without costs. *Journal of experimental psychology. Human perception and performance*, *44* (3), 398–411. doi: 10.1037/xhp0000449

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology*, *102* (4), 726–747. doi:10.1111/j.2044-8295.2011.02053.x

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86* (3), 214–255. doi: 10.1037/0033-295X.86.3.214

Nee, D. E., & Jonides, J. (2013). Trisecting representational states in short-term memory. *Frontiers in human neuroscience*, *7*, 796. Doi: 10.3389/fnhum.2013.00796

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661-677.

New, B., Pallier, C., Brysbaert, M., Ferrand, L. (2004) Lexique 2 : A New French Lexical Database. *Behavior Research Methods, Instruments, & Computers*, *36* (3), 516-524. doi: 10.3758/bf03195598

Nieuwenstein, M., & Wyble, B. (2014). Beyond a mask and against the bottleneck: retroactive dual-task interference during working memory consolidation of a masked visual target. *Journal of experimental psychology. General*, *143* (3), 1409–1427. doi: 10.1037/a0035257

Nishiyama, R. (2018). Separability of active semantic and phonological maintenance in verbal working memory. *PloS One*, *13*, e0193808.

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411-421.

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 45–100). Elsevier Academic Press. doi: 10.1016/S0079-7421(09)51002-X

Oberauer, K. (2019). Working Memory and Attention: A Conceptual Analysis and Review. *Journal of Cognition*, *2(1)*, 36. doi: 10.5334/joc.58

Oberauer, K., & Hein, L. (2012). Attention to information in Working Memory. *Current Directions in Psychological Science*, *21 (3)*, 164-169.

Oberauer, K., Lewandowsky, F., Farrell, W., Jarrold, C. & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*, 779-819. doi: 10.3758/s13423-012-0272-4

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological bulletin*, *144 (9)*, 885–958. doi: 10.1037/bul0000153

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*, 195-203.

Peterson, L. R.; Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*. *58 (3)*, 193–198. doi:10.1037/h0049234.

Pillon, B., Bazin, B., Deweer, B., Ehrlé, N., Baulac, M., & Dubois, B. (1999). Specificity of Memory Deficits after Right or Left Temporal Lobectomy. *Cortex*, *35 (4)*, 561-571

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly journal of experimental psychology*, *48* (2), 384–404. doi: 10.1080/14640749508401396

Poltrock, S. E., Lansman, M., & Hunt, E. (1982). Automatic and controlled attention processes in auditory target detection. *Journal of experimental psychology. Human perception and performance*, *8* (1), 37–45. doi: 10.1037//0096-1523.8.1.37

Portrat, S., & Lemaire, B. (2015). Is attentional refreshing in working memory sequential? A computational modelling approach. *Cognitive Computation*, *7* (3), 333–345. Doi: 10.1007/s12559-014-9294-8

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*. *32* (1), 3–25.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

Rademaker, R. L., Tredway, C. H., & Tong, F. (2012). Introspective judgments predict the precision and likelihood of successful maintenance of visual working memory. *Journal of Vision*, *12*, 1–13.

Ranganath, C. C., Michael, B. X., & Craig, J. B. (2005). Working Memory Maintenance Contributes to Long-term Memory Formation: Neural and Behavioral Evidence. *Journal of Cognitive Neuroscience*, *17* (7), 994–1010.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A Minimal Executive Function. *Cortex*, *43*, 135-145.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A., & Greene, E. J. (2002). Neuroimaging a Single Thought: Dorsolateral PFC Activity Associated with Refreshing Just-Activated Information. *NeuroImage*, *15*, 447-53

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18* (3), 849–60. doi:10.1037/0096-1523.18.3.849

Rey, A. E., Versace, R., & Plancher, G. (2018). When a Reactivated Visual Mask Disrupts Serial Recall: Evidence That Refreshing Relies on Memory Traces Reactivation in Working Memory. *Experimental Psychology*, *65*, 1-9.

Rhodes, S., Jaroslawska, A. J., Doherty, J. M., Belletier, C., Naveh-Benjamin, M., Cowan, N., Camos, V., Barrouillet, P., & Logie, R. H. (2019). Storage and processing in working memory: Assessing dual task performance and task prioritization across the adult lifespan. *Journal of Experimental Psychology: General*, *148*, 1204-1227.

Richmond, L. L., Burnett, L. K., Morrison, A. B., & Ball, B. H. (2021). Performance on the processing portion of complex working memory span tasks is related to working memory capacity estimates. *Behavior Research Method*, *53* (5)

Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, *2*, 259–279.

Ricker, T. J., & Cowan, N. (2010). Loss of Visual Working Memory Within Seconds: The Combined Use of Refreshable and Non-Refreshable Features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36* (6), 1355–1368

Ricker, T. J., & Cowan, N. (2014). Differences between presentation methods in working memory procedures: a matter of working memory consolidation. *Journal of experimental psychology. Learning, memory, and cognition*, *40* (2), 417–428. doi: 10.1037/a0034301

Ricker, T. J., & Hardmann, K. O. (2017). The Nature of Short-Term Consolidation in Visual Working Memory. *Journal of Experimental Psychology: General*, *146* (11), 1551–1573.

Ricker, T. J., & Vergauwe, E. (2020). Consistent failure to produce a cognitive load effect in visual working memory using a standard dual-task procedure. *Journal of Cognition*, 3(1), 1–10. doi: 10.5334/joc.108

Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory. *Annals of the New York Academy of Sciences*, 1424 (1), 8–18. doi: 10.1111/nyas.13633

Ricker, T. J., Vergauwe, E., & Cowan, N. (2014). Decay Theory of Immediate Memory: From Brown (1958) to Today (2014). *Quarterly Journal of Experimental Psychology*, 69, 1969–1995.

Rubenstein, H., Garfield, L., & Amilukan, J. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, 9, 487-494.

Saint-Aubin, J., Ouellette, D., & Poirier, M. (2005). Semantic similarity and immediate serial recall: is there an effect on all trials? *Psychonomic bulletin & review*, 12 (1), 171–177. doi: 10.3758/bf03196364

Scarborough, D., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 1–17.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.

Schurgin, M. V., Cunningham, C. A., Egeth, H. E., & Brady, T.F. (2018). Visual Long-term Memory Can Replace Active Maintenance in Visual Working Memory. *Biorxiv*. doi: 10.1101/381848

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory and Cognition*, 21, 168-175.

Shapiro, K. (1997). The Attentional Blink. *Trends in Cognitive Sciences*. 1 (8), 291–296. doi:10.1016/S1364-6613(97)01094-2

Shivde, G., & Anderson, M. C. (2011). On the existence of semantic working memory: Evidence for direct semantic maintenance. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 37, 1342–1370.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6 (2), 174–215. Doi:10.1037/0278-7393.6.2.174

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215. Doi:10.1037/0278-7393.6.2.174

Soldan, A., Zarah, E., Hilton, H. J. & Stern, Y. (2008). Global familiarity of visual stimuli affects repetition-related neural. *Neuroimage*, 39 (1), 515–526.

Souza, A. & Oberauer, K. (2018). Does articulatory rehearsal help immediate serial recall? *Cognitive Psychology*, 107, 1-21.

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, 1339, 20–31. doi: 10.1111/nyas.12603

Souza, A. S., Vergauwe, E., & Oberauer, K. (2018). Where to attend next: guiding refreshing of visual, spatial, and verbal representations in working memory. *Annals of the New York Academy of Sciences*, 1424 (1), 76–90. doi:10.1111/nyas.13621

Souza, A., & Oberauer, K. (2016). In search of the focus of attention in WM: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, 78, 1839-60.

Souza, A., & Oberauer, K. (2017). The contributions of visual and central attention to visual working memory. *Attention, Perception & Psychophysics*, *79*, 1897-1916.

Souza, A., Overkott, C., & Matyja, M. (2021). Categorical distinctiveness constrains the labeling benefit in visual working memory. *Journal of Memory and Language*, *119*, doi: 10.1016/j.jml.2021.104242

Souza, A., & Skóra, Z. (2017). The interplay of language and visual perception in working memory. *Cognition*, *166*, 277-297.

Souza, A., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual WM. *Annals of the New-York Academy of Sciences*, *1339*, 20-51.

Souza, A., Rerko, L., Lin, H-Y., & Oberauer, K. (2014). Focused Attention improves WM: implication for flexible-resource and discrete capacity models. *Attention, Perception, & Psychophysics*, *76*, 2080-2102.

Souza, A., Vergauwe, E., & Oberauer, K. (2018). Where to attend next? Guiding refreshing of visual, spatial and verbal representations in WM. *Annals of the New-York Academy of Sciences*, *1424*, 76-90.

Souza-Talarico, J. N. de, Caramelli, P., Nitri, R., & Chaves, E. C. (2007). The influence of schooling on working memory performance in elderly individuals without cognitive decline. *Dementia & Neuropsychologia*, *1* (3), 276–281. doi:10.1590/s1980-57642008dn10300009

Sperber, R. D., McCauley, C., Ragain, R. D., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory & Cognition*, *7* (5), 339–345. doi:10.3758/bf03196937

Sperling, G., & Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science*, *202*, 315–318.

Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: an event-related potential study. *Neuroreport*, *19* (5), 583–587. doi: 10.1097/WNR.0b013e3282f97b4d

Stelter, M., & Degner, J. (2018). Investigating the other-race effect in working memory. *British journal of psychology*, *109* (4), 777–798. doi: 10.1111/bjop.12304

Stevanovski, B. & Jolicoeur, P. (2007). Visual short-term memory: central capacity limitations in short-term consolidation. *Visual Cognition*, *15*, 532–563.

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Science*. *19* (7), 394-405. doi: 10.1016/j.tics.2015.05.004

Sun, H., Zimmer, H. D., & Fu, X. (2011). The influence of expertise and of physical complexity on visual short-term memory consolidation. *Quarterly Journal of Experimental Psychology*, *64*, 707–729.

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *The Quarterly Journal of Experimental Psychology*, *62* (3), 435–443. doi: 10.1080/17470210802386375

Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review*, *15*, 535-542.

Tanaka, J. W., Heptonstall, B., & Hagen, S. (2013) Perceptual expertise and the plasticity of other-race face recognition. *Visual Cognition*, *21* (9-10), 1183-1201. doi: 10.1080/13506285.2013.826315

Thalman, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45* (1), 37-55.

Thorn, A.S.C., Gathercole, S.E. & Frankish, C.R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert's (1993) multinomial processing tree model. *Cognitive Psychology*, *50* (2), 133–158. doi: 10.1016/j.cogpsych.2004.07.001

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language*, *39* (2), 195–217. doi: 10.1006/jmla.1998.2574

Towse, J. N., Hitch, G. J., & Hutton, U. (2002). On the nature of the relationship between processing activity and item retention in children. *Journal of Experimental Child Psychology*, *82* (2), 156–184. doi: 10.1016/S0022-0965(02)00003-6

Towse, J., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *48A* (1), 108–124. doi: 10.1080/14640749508401379

Tripathy, S. P., & Öğmen, H. (2018). Sensory Memory Is Allocated Exclusively to the Current Event-Segment. *Frontiers in psychology*, *9*, 1435. Doi: 10.3389/fpsyg.2018.01435

Tsubomi, H., Fukuda, K., Watanabe, K., & Vogel, E. K. (2013). Neural limits to representing objects still within view. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *33* (19), 8257–8263. doi: 10.1523/JNEUROSCI.5348-12.2013

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*, *114* (1), 104–132. doi: 10.1037/0033-295X.114.1.104

Valentine, T., & Endo, M. (1992). Towards an Exemplar Model of Face Processing: The Effects of Race and Distinctiveness. *The Quarterly Journal of Experimental Psychology Section A*, *44* (4), 671–703. doi:10.1080/14640749208401305

Vergauwe, E. (2018). Comparing different instructed-refreshing schedules: evidence for cumulative, forward-order refreshing of verbal lists? *Annals of the New York Academy of Sciences*, *1424*, 102-114.

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience*, *8*, Article 126. doi: 10.3389/fnhum.2014.00126

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: Evidence that refreshing and memory search are closely related. *Psychonomic Bulletin and Review*, *22*, 1001-1006.

Vergauwe, E., & Langerock, N. (2017). Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *Journal of Memory and Language*, *96*, 23-35.

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35* (4), 1012–1028. doi:10.1037/a0015859

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological science*, *21* (3), 384–390. doi: 10.1177/0956797610361340

Vergauwe, E., Camos, V. & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *40* (4). 1072-1094. Doi: 10.1037/a0035779

Vergauwe, E., Dewaele, N., Langerock, N., & Barrouillet, P. (2012). Evidence for a central pool of general resources in working memory. *Journal of Cognitive Psychology*, 24 (3), 359–366. doi: 10.1080/20445911.2011.640625

Vergauwe, E., Hardman, K. O., Rouder, J. N., Roemer, E., McAllaster, S., & Cowan, N. (2016). Searching for serial refreshing in working memory: Using response times to track the content of the focus of attention over time. *Psychonomic Bulletin & Review*, 23 (6), 1818–1824. doi: 10.3758/s13423-016-1038-1

Vogel, E.K., S.J. Luck & K.L. Shapiro. (1998). Electrophysiological evidence for a post perceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*. 24, 1656–1674.

Vogel, E.K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 1436–1451.

Voyer, D. (2003). Word frequency and laterality effects in lexical decision: Right hemisphere mechanisms. *Brain & Language*, 87 (3), 421-431.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science*, 238 (4828), 778–780. doi:10.1126/science.3672124

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*. 142 (5), 472–97. doi:10.1037/bul0000030

Wood, R., Baxter, P.; & Belpaeme, T. (2011). A review of long-term memory in natural and synthetic systems. *Adaptive Behavior*, 20 (2), 81–103. doi:10.1177/1059712311421219

Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 787–807.

Wyble, B., M.C. Potter, H. Bowman, et al. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, *140*, 488–505.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological research*, *63* (3-4), 212–233. doi: 10.1007/s004269900003

Xie, W., & Zhang, W. (2018). Familiarity Speeds Up Visual Short-term Memory Consolidation: Electrophysiological Evidence from Contralateral Delay Activities. *Journal of cognitive neuroscience*, *30* (1), 1–13. doi: 10.1162/jocn_a_01188

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81* (1), 141–145.

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

Zhou, X., Mondloch, C. J., & Emrich, S. M. (2018). Encoding differences affect the number and precision of own-race versus other-race faces stored in visual working memory. *Attention, perception & psychophysics*, *80* (3), 702–712. doi: 10.3758/s13414-017-1467-6