A Time-Based Resource-Sharing Account of Switching Costs between: Processing and Storage in Working Memory

Miriam Debraise

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE SOCIÉTÉS, HUMANITÉS, ARTS ET LETTRES

THÈSE DE DOCTORAT

Coûts d'alternance entre traitement et maintien d'informations en mémoire de travail : Mise à l'épreuve du modèle de partage temporel des ressources

Miriam DEBRAISE

Laboratoire BCL : Bases, Corpus, Langage - UMR 7320 CNRS

Présentée en vue de l'obtention du grade de docteur en Psychologie
Université Côte d'Azur
Dirigée par : Fabien Mathy
Co-encadrée par : Nicolas Gauvrit
Soutenue le : 25 juin 2021

Devant le jury, composé de :
Valérie CAMOS, Professeure, Université de Fribourg
Nicolas GAUVRIT, MCF-HDR, Université de Lille
Fabien MATHY, Professeur, Université Côte d'Azur
Sophie PORTRAT, MCF, Université Grenoble Alpes
Arnaud SZMALEC, Professeur, Université catholique de Louvain
Evie VERGAUWE, Professeure associée, Université de Genève

Coûts d'alternance entre traitement et maintien d'informations en mémoire de travail : Mise à l'épreuve du modèle de partage temporel des ressources

Directeurs de thèse:
Nicolas Gauvrit, MCF-HDR, Université de Lille
Fabien Mathy, Professeur, Université Côte d'Azur

Président du jury:
Valérie Camos, Professeure, Université de Fribourg

Rapporteurs:
Sophie Portrat, MCF, Université Grenoble Alpes
Arnaud Szmalec, Professeur, Université catholique de Louvain

Examinateur:
Evie Vergauwe, Professeure associée Université de Genève

A Time-Based Resource-Sharing Account of Switching Costs between Processing and Storage in Working Memory

PhD Advisors:
Nicolas Gauvrit, Associate Professor, Université de Lille
Fabien Mathy, Full Professor, Université Côte d'Azur

President of the jury:
Valérie Camos, Full Professor, Université de Fribourg

Referees:
Sophie Portrat, Associate Professor, Université Grenoble Alpes
Arnaud Szmalec, Full Professor, Université catholique de Louvain

Examinator:
Evie Vergauwe, Associate Professor, University of Geneva

Abstract

Working memory allows the temporary storage and processing of a limited amount of information. According to the Time-Based Resource-Sharing Model, working memory functions by rapidly switching between storage and processing activities. The complex span tasks generally used to evaluate this process require this rapid alternation. The task involves a primary memory task interspersed with a secondary processing task designed to direct attention away from the memory task.

The present dissertation aimed at manipulating the structure of the complex span task to examine untested predictions from the TBRS. I first identified in the literature two types of structures across the available complex span tasks. The unnoticed difference (at least conceptually) between the two types of tasks relies on whether the task begins with a processing episode (that captures attention) or by a storage episode (that requires the encoding of a to-be-recalled item). After providing evidence that this discrepancy does have an influence on the estimate of working memory capacity, I pursue my exploration of more subtle induced variations of task structure on working memory performance.

I argue that how storage and processing activities are distributed within a concurrent task may induce switching costs in working memory performance that have gone unnoticed in past and current experimental studies.

Switching costs are induced by any kind of multitasking situations that complex span tasks should not be exempt of. We tested whether switching costs could be implemented into the TBRS to improve fit to the data. For this purpose, we used complex span tasks that involved elementary processing steps presented at a predefined pace to manipulate switch costs. The number of switches was manipulated by the use of different rhythmic patterns within the concurrent task. The results showed that these patterns can slightly influence working memory performance in some cases but not in others. However, the rhythmic patterns used in our secondary tasks may have introduced a confound variable as some of the patterns containing fewer alternations between storage and processing also introduced longer delays of free time, hence possibly enhancing the consolidation of the memoranda. A last part of the thesis addresses this confound using rhythmic patterns designed to test specifically the switching cost and the consolidation accounts. We conclude that the TBRS in its original form offers a satisfying fit of the data.

Keywords: Working memory, complex span task, switching costs, TBRS model

Résumé

La mémoire de travail permet le stockage et le traitement temporaire d'une quantité limitée d'informations. Selon le modèle du partage temporel des ressources (Time-Based Resource-Sharing model - TBRS), la mémoire de travail fonctionne grâce à une alternance rapide entre le stockage et le traitement. Les tâches d'empan complexes généralement utilisées pour évaluer ce double processus utilisent une tâche de mémoire entrecoupée par une tâche secondaire. La tâche secondaire est conçue pour détourner l'attention. L'objectif principal du présent travail était de manipuler la structure de la tâche d'empan complexe afin d'éprouver le modèle TBRS.

Nous avons identifié en premier lieu dans la littérature deux types de structures de tâche d'empan complexe. Cette différence négligée jusqu'alors (du moins conceptuellement) entre ces deux types de tâches repose sur l'ordre des activités de stockage et de traitement. La tâche commence soit par un épisode de traitement (qui capture l'attention) soit par un épisode de stockage (qui nécessite l'encodage d'un élément à rappeler). Après avoir montré que cette divergence de procédure influence les mesures de capacité en mémoire de travail, nous poursuivons notre exploration en construisant des variations plus subtiles de la structure de la tâche concurrente. L'idée est que l'ordre des activités de traitement et de maintien au sein d'une tâche concurrente pourrait induire des coûts de switching en mémoire de travail. Nous faisons l'hypothèse que ces coûts de switching ne devraient pas faire exception dans les tâches d'empan complexe, puisqu'ils sont présents dans d'autres activités nécéssitant d'alterner entre plusieurs tâches. Nous avons donc proposé d'étudier si ces coûts d'alternance pouvaient être implémentés dans le TBRS afin d'améliorer ses prédictions. Dans cet objectif, nous avons utilisé des tâches d'empan complexes impliquant des épisodes de traitements élémentaires présentées à un rythme prédéfini. Le nombre de switching a été manipulé par l'utilisation de différents patrons rythmiques au sein de la tâche concurrente. Les résultats ont montré que ces coûts d'alternance peuvent légèrement influencer les performances en mémoire de travail, mais de façon ni franche ni systématique. Néanmoins, les patrons rythmiques utilisés dans nos tâches secondaires pourraient avoir introduit une variable confondue. En effet, certains patrons rythmiques ont introduit des délais plus longs de temps libres susceptibles d'améliorer la consolidation en mémoire. Une dernière partie de cette thèse aborde cette question afin de dissocier les potentiels effets de consolidation et de switching. Nous concluons que le modèle TBRS permet de rendre compte des données expérimentales de façon satisfaisante.

Mots clés: Mémoire de travail, tâche d'empan complexe, coûts d'alternance, modèle TBRS

Acknowledgement

Avant tout, je tiens à remercier les membres du Jury, Valérie Camos, Sophie Portrat, Arnaud Szmalec et Evie Vergauwe de m'avoir fait l'honneur de participer à l'évaluation de ce travail.

Je remercie très chaleureusement mes directeurs de thèse pour leur soutien et leur confiance. Merci à Fabien Mathy pour ses conseils toujours excellents, sa disponibilité plus qu'impressionnante, son pragmatisme et pour m'avoir guidée de manière si constructive. Merci à Nicolas Gauvrit pour sa disponibilité malgré la distance géographique et les échanges instructifs sur le modèle TBRS2.

Merci à Sophie Portrat et à Benoit Lemaire pour les suggestions, les conseils, les corrections lors de mes premiers posters, ainsi que pour leur insight sur le modèle TBRS*.

Merci à la direction et à l'ensemble des membres de BCL pour leur accueil chaleureux et la bonne ambiance, mais aussi pour les financements qui m'ont permis de participer aux conférences.

Merci Anja Dordevic, Mélanie Laine et Charlotte Lang qui ont contribué aux passations expérimentales. Merci à Christophe Cauchi pour avoir accepté de partager ses données. J'adresse aussi mes remerciements aux 740 étudiants ayant participé aux expériences.

Suuri kiitos perheelleni ja läheisilleni. Uskoitte että pystyn tähän, silloinkin, kun minä en siihen uskonut. Erityisesti kiitos Äiti et Papa ehtymättomästä tuesta ja rakkaudesta. Merci à ma sœur Nora de toujours tout comprendre (♡).

Merci à mes amis pour leur précieux soutien...Merci aux doctorants et docteurs de la MSHS pour leur multiples conseils et tous les moments de joie partagés...Rakkaille ystävilleni kiitos, olette korvaamaton iso osa elämääni etäisyydestä huolimatta. Kiitos että tuotte pirskahtelevaa ja kuplivaa iloa harmaissakin hetkissä...

Enfin, pour terminer sur le meilleur, merci à Cyril pour son soutien sans faille, pour être tout simplement là et de toujours trouver les mots pour me rassurer dans les moments de doutes ou me faire rire dans les moments désespoirs. Merci aussi pour son aide et sa patience infinie pour les cours de programmation ou de math... ce fût beaucoup de travail (comme pour un album d'Astérix). Cette thèse c'est aussi la tienne.

Contents

Outline

This PhD dissertation studies working memory, which refers to a complex cognitive function that allows the temporary storage and processing of a limited amount of information. I examine variants of the Time-Based Resource-Sharing (TBRS) model, that considers that working memory functions by rapidly switching between processing and storage activities. The experimental method is based on the complex span task. This task requires a rapid alternation between a memory task and a secondary processing task. The secondary task is designed to direct attention away from the memory task.

The main goal of the present dissertation was to manipulate the temporal structure of the complex span task to explore untested predictions from the TBRS.

Chapter 1 provides a general background of the existing studies on working memory and experiments that have focused on the complex span task. This task has been explicitly designed to tap both storage and processing activities in working memory.

In **Chapter 2**, I point out that sequential structure of the complex span task varies from one study to another. Some studies alternate storage and processing activities, while others alternate inversely processing and storage activities. The question is whether these variants could influence estimates of working memory capacity. For this purpose, we ran an experiment comparing these two structures of the complex span task. The result shows that the choice of task does influence working memory measures. This finding encouraged us to pursue in the following chapters our investigation on the potential influence of more refined variations of the processing activities on working memory capacity.

Chapter 3 explores further the structure of the processing task at a finer level. More specifically, this chapter focuses on the influence of switching costs caused by the sequential structure of the task. Switching costs are induced by multitasking and our objective was to suggest that they could be better implemented into the TBRS model. For this purpose, we ran a series of experiments using novel complex span tasks in which the number of switches was manipulated with different rhythmic patterns. The results show that these patterns can slightly influence working memory performance in some cases.

In **Chapters 4 and 5**, I address the issue that our manipulation of the switches may have introduced a confound as some of the rhythmic patterns introduce longer delays of free time, possibly enhancing consolidation. We explored this confound with another series of experiments. The results showed some evidence supporting switching costs in working memory.

The Conclusion in **Chapter 6** draws together the observations arising from the experiments analyzed in the previous chapters. We conclude that the original TBRS model is mostly satisfying and parsimonious enough to account for a large range of structural modifications of the complex span task.

❧

1 Introduction

The construct of working memory refers to our ability to maintain and process information temporarily. This ability is central to most of our daily life activities. For instance, reading the following text requires working memory, as the reader needs to keep in mind earlier processed information in order to keep track of the whole meaning of this text. Working memory is indeed essential for (reading) comprehension but also for other complex cognitive activities such as problem-solving, planning, and reasoning which are central to learning processes (Cowan, 2014), and which all require the manipulation of information on the spot.

The term *working memory* was coined by Miller, Galanter, and Pribram (1960) to describe a mental place able to hold and monitor our intentions and goals[1]. More than a half a century from its first apparition, working memory has become one of the most popular topics of research within the field of cognitive science (Miyake & Shah, 1999). In point of fact, as I write this introduction, a simple Google Scholar search request containing the term *working memory* in the title of an article has more than 56.000 hits, and by the end of the year 2021 this number is expected to increase by more than a thousand if the trend of the last decay continues. The myriad of terminology and definitions associated to the object of study in question can only substantiate this phenomenon.

As discussed above, the keen interest in working memory, and in memory in general, has long fascinated investigators. The first scientific study on memory is often attributed to Hermann Ebbinghaus at the end of the 19th century, who explored his own learning capacities and forgetting curve of newly acquired information over a range of different time intervals. Ebbinghaus (1885/1913) describes moments of "especial mental clearness" where information is almost memorized. He also adds that these moments can be compared to a "first fleeting grasp" of the information to be memorized, with no guarantee that it will be successfully recalled later on. At the same time, James (1890) proposed the notions of *primary memory* and secondary memory. Primary memory was defined as the limited information held in the conscious present, as opposed to the secondary memory which included

[1] Although, according to Adams, Nguyen, and Cowan (2018), the term had actually already been used prior to Miller et al. (1960) in the field of computer science.

information held indefinitely outside of consciousness that could be brought back, if desired (James, 1890). According to Cowan (2008), this primary memory can be seen as akin to this phenomenon of fleeting grasp first described by Ebbinghaus.

Later, the term *short-term memory* or *short-term storage* was popularized by several authors (e.g. Atkinson & Shiffrin, 1968; Broadbent, 1958). Short-term memory was described as a temporary and limited capacity memory, the only difference with the concept of primary memory being that not everything in short-term memory is necessarily in a state of conscious awareness (Cowan, 2008).

Conceptually, short-term memory is very close to working memory. The distinction often made between the two notions is that short-term memory is the ability to maintain passively in a highly activated state a small amount of information for a short period of time, whereas working memory adds an active dimension to it, entailing thus both storage *and* processing of information. However, there is no consensus about the use of the terms or the exact nature of the relation between the two constructs. For instance, according to Cowan (2008, p. 3), "(...) working memory includes short-term memory and other processing mechanisms that help to make use of short-term memory. This definition is different from the one used by some other researchers (e.g. Engle, 2002), who would like to reserve the term working memory to refer only to the attention-related aspects of short-term memory. This, however, is not so much a debate about substance but rather a slightly confusing discrepancy in the usage of terms."

Aside the diverse terminology revolving around the notion of working memory (e.g. immediate memory, short-term memory), Cowan (2017) delineate a whopping amount of nine different definitions of working memory frequently used by investigators, as seen in Figure 1.1. The diversity of working memory definitions has undoubtedly created some confusion among researchers making it hard to sometimes share theoretical positions. While two theories may share the same definition of working memory, it is also common that the definition of working memory is modified to fit the assumptions of a novel theory (Cowan, 2017). This seems very peculiar, as we can not imagine different chemists working with different periodic tables.

As seen in Figure 1.1, it is difficult to find some common ground between the definitions presented, since so many of them seem driven by incompatible theoretical assumptions (each favored by different sets of empirical facts that most

1. Computer WM (e.g., Laird, 2012; Newell & Simon, 1956)
A holding place for information to be used temporarily, with the possibility of many working memories being held concurrently.

2. Life-planning WM (e.g., Miller et al., 1960)
A part of the mind that saves information about goals and subgoals needed to carry out ecologically useful actions.

3. Multicomponent WM (e.g., Baddeley, 1986, 2000; Baddeley & Hitch, 1974)
A multicomponent system that holds information temporarily and mediates its use in ongoing mental activities.

4. Recent-event WM (e.g., Olton et al., 1977)
A part of the mind that can be used to keep track of recent actions and their consequences in order to allow sequences of behaviors to remain effective over time.

5. Storage-and-processing WM (e.g., Daneman & Carpenter, 1980)
A combination of temporary storage and the processing that acts upon it, with a limited capacity for the sum of storage and processing activities. When the storage component alone is measured, or the processing component alone is measured, the term WM is not applied, in contrast to the usage within multicomponent WM. Further distinguishing this definition from multicomponent WM, there is not always a clear commitment to multiple storage components, only a separation between storage and processing.

6. Generic WM (e.g., Cowan, 1988)
The ensemble of components of the mind that hold a limited amount of information temporarily in a heightened state of availability for use in ongoing information processing.

7. Long-term WM (e.g., Ericsson & Kintsch, 1995)
The use of cue and data-structure formation in long-term memory that allows the information related to an activity to be retrieved relatively easily after a delay.

8. Attention-control WM (e.g., Engle, 2002)
The use of attention to preserve information about goals and sub-goals for ongoing processing and to inhibit distractions from those goals; it operates in conjunction with short-term storage mechanisms that hold task-relevant information in a manner that does not require attention.

9. Inclusive WM (e.g., Unsworth & Engle, 2007)
The mental mechanisms that are needed to carry out a complex span task; it can include both temporary storage and long-term memory, insofar as both of them require attention for the mediation of performance.

Figure 1.1: The nine definition of working memory (WM) according to Cowan (2017). Reprinted from "The many faces of working memory and short-term storage", by N. Cowan, 2017, Psychonomic bulletin review, 24(4), p. 1159 (https://doi.org/10.3758/s13423-016-1191-6). Copyright © 2016 by Psychonomic Society, Inc

often hold in a restricted experimental paradigm). However, as argued by Cowan (2017) "It seems important to strive for definitions general enough that an empirical investigation is unlikely to overturn the definition itself any time soon.". For this reason, we use in the present work the generic definition of working memory (Cf. Generic WM in Figure 1.1), which defines working memory as "The ensemble of components of the mind that hold a limited amount of information in heightened state of availability for use in ongoing information processing" (Cowan, 2017, p. 1163).

Perhaps one of the reasons why working memory has been such a fruitful avenue of research is the predictive power of working memory tasks. Several studies have indeed shown that working memory tasks are highly predictive of higher order cognition such as comprehension, learning and reasoning abilities (e.g. Conway, Kane, & Engle, 2003; Daneman & Merikle, 1996; Engle, Kane, & Tuholski, 1999; Kyllonen & Christal, 1990; Kane, Hambrick, & Conway, 2005; Unsworth & Engle, 2007b). Because intelligence too has the power to predict educational and occupational outcomes decades later (Lubinski, 2004), the extrapolation that working memory could lie at the root of an individual's success is tempting. In the next section, we will discuss the most commonly used tasks of short-term memory, working memory and the capacity limits that have been inferred from these tasks.

1.1 Measures of working memory

The most canonical task to measure short-term memory capacity is the span task. It was used in the first intelligence test devised by Binet and it has been implemented since in the most widely used Wechsler's intelligence tests (see Ackerman, Beier, & Boyle, 2005 for a review). Performance at this task has been constant for the last 100 years, so it does not seem sensitive to education (Gignac, 2015).

Typically, verbal and visual simple span tasks are measured by the digit span and the Corsi blocks respectively. In these tasks, participants are required to remember in the correct serial order a series of items. However, simple span tasks were thought to measure only a passive storage ability. Therefore, the more active dimension of working memory required to design more adequate tasks. The complex span task was explicitly designed for this purpose, evaluating thus both storage and processing activities. This transition in the experimental procedures aiming at estimating the capacity of immediate memory[2] facilitated the conceptual transition from the passive short-term memory construct to the more active working memory construct.

The inaugural complex span task, the reading span task, was developed by Daneman and Carpenter (1980). In this task, participants are asked to read unrelated sentences, while recalling only the last word of each sentence. The complex span task quickly became one of the most popular measures of working memory with several existing variants. Such variants include for instance the operation span (Turner & Engle, 1989) and the symmetry span (Shah & Miyake, 1996). More recently, the parity judgment span and the reading digit span task have been designed to involve elementary processing steps (Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007).

In each of these variants, items (e.g. letters, words, images, digits) are to be recalled while a secondary task is executed at the same time. For instance, in the operation span, participants are required to solve arithmetic calculations while memorizing consonants. In the symmetry span, the secondary task is a 8×8 grid in which cells are filled in black to form a pattern. The participants are required to evaluate if the pattern is symmetrical on the vertical axis or not, while memorizing individual grid locations given serially. The reading digit and the parity judgment task are similar visually but the nature of the task at hand differs between the two. In both tasks, the to-be-recalled letters are presented serially and interspersed with a display showing a digit. Participants are either asked to read aloud the digits or they are required to evaluate whether the digit in between two letters is even or odd. In both cases, the final memory test consist in recalling the letters in their correct order. Figure 1.2 illustrates these four complex span tasks, which again are just a few examples among many possible variants. Although the to-be-recalled items and the secondary task vary, the dual methodology of interleaving storage and processing subtasks is common to all complex span tasks.

1.1.1 Limits in defining the limit

According to Baddeley, Thomson, and Buchanan (1975), short-term memory capacity is limited by the amount of time during which the memorandum can remain

[2] Immediate memory can be used as an umbrella term when presenting short-term memory and working memory indistinctively in a historical presentation.

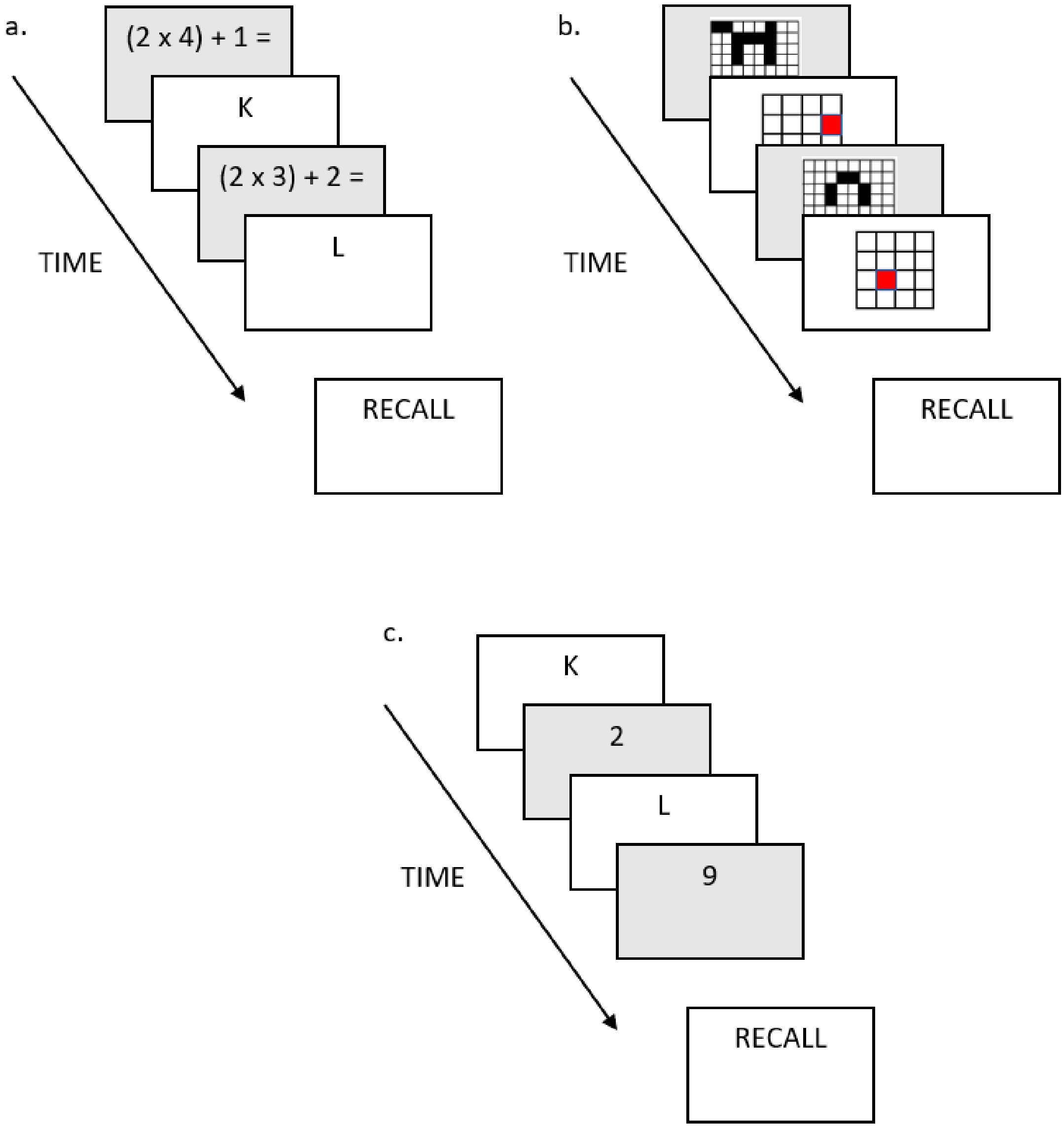


Figure 1.2: Examples of the Operation span, the Symmetry span, the Reading Digit, and the Parity Judgment span tasks. Figure a. represents the Operation span. Figure b. represents the Symmetry span. Figure c. illustrates the Reading Digit or the Parity Judgment task. The white *vs.* grey displays represents the storage items and the secondary task respectively. Adapted from "Shortened complex span tasks can reliably measure working memory capacity", by J.L. Foster et al., (2015), Memory Cognition, 43(2), 226 (https://doi.org/10.3758/s13421-014-0461-7). Copyright © 2014, Psychonomic Society, Inc.

activated using rehearsal strategies (i.e. overt or covert articulatory repetition of memoranda). This limit has been estimated to be about 1.8 seconds (see Baddeley, 1986). The combination of this time limit and speech rate determines the individual capacity limit in terms of verbal material which can be retained thanks to rehearsal (Hulme, Thomson, Muir, & Lawrence, 1984).

Despite the value of this attempt to define capacity limit in terms of a continuous duration, a more straightforward characterization of working memory capacity has been the span because it is a discrete measure. Whether it is calculated from a simple span task or a complex span task, the span corresponds to the maximal number of items that a person can hold in view of an immediate recall (Conway et al., 2005).

One of the earliest studies on memory span was conducted by Jacobs (1887). Jacob's early work had identified that participants have a superior span when the recalled material were digits as compared to letters, which were in turn memorized more successfully than nonsense syllables stimuli. But still, performance revolved around 7, with the average span estimated around 9, 7, and 6 for digits, letters and the nonsense syllables respectively. This is in line with the infamous article of Miller (1956), which noted that several experiments from the 1950's converged towards –what he called the magical number– 7 as a general estimate of short-term memory capacity.

Since then, several authors have suggested that Miller's magical number is an overestimation of the true capacity. The span of 7 is thought to reflect short-term memory capacity as measured by simple span tasks. However, individuals are far from being passive in these tasks as mnemonics strategies can be used freely by participants to enhance their recall performance. When they cannot, as in the complex span tasks in which the second task hinders strategies such as rehearsal, the pure storage capacity has been best estimated to four slots on average (Broadbent, 1975; Cowan, 2001; Halford, Mayberry, & Bain, 1988; Halford, Wilson, & Phillips, 1998; Luck & Vogel, 1997). In the simplest span tasks, when the span can be increased to 7 items, individuals artificially group information by using verbal strategies. The amount of information retained still corresponds to the true capacity, but it can be considered that the four slots of information have been filled up with larger chunks of several items (Mathy & Feldman, 2012).

Regarding the limitations of the span, one can only speculate about why the span corresponds to a certain number, like 4 (as it seems to be the case since many different types of working memory tasks converge towards this limit; see Cowan, 2001). Although combinatorial issues seem to start with four dimensions (Broadbent, 1975; Halford, Baker, McCredden, & Bain, 2005), there is a priori no cognitive reasons for why humans do not have a larger capacity, and attempts to account for why capacity is limited to 4 items can quickly lead to a dead-end.

A more practical approach to study capacity limits has been to focus on predicting within- and between-subjects variations of the span (e.g. Kane et al., 2004, for instance in correlational studies (e.g. Unsworth & Engle, 2007a). However, correlational studies have their own limitations with regard to identifying the mechanisms at play in capacity limits, because there is no way to manipulate factors in this method.

The approach taken in the present dissertation has been to design canonical experimental tasks to study variations of working memory capacity, with the advantage that experimental designs are conducive to computational approaches.

1.2 Models of working memory

Over the years, numerous theories and models of working memory have been proposed. This prolific landscape points out the complexity of working memory research. One key aspect where different models diverge is how they explain the (main) source of forgetting in working memory. Two alternative hypothesis are put forward to explain forgetting in working memory: the decay or the interference account (Oberauer, Farrell, Jarrold, & Lewandowsky, 2016).

According to decay-based models, forgetting occurs because memory traces suffer from a progressive loss of activation, called decay, with the passage of time. Simply, disuse of memory is seen as the cause of forgetting (Ricker, Vergauwe, & Cowan, 2016). Although this account seems intuitive, it has been the subject of much criticism since its inception. As noted by the authors, despite the fact that the issue has been debated for over a half a century, the arguments against decay have not much evolved, and a frequent comment is that time only gives other mechanisms than decay an opportunity to operate. To make this point, the following classic analogy made by McGeoch (1932) is often reported: rust accumulates on metal over time, however, time itself is not responsible for this effect. Rather, the true culprit is oxidation which is indeed a time-dependent process. The same argument is used against decay: the sole role of misuse of the memory traces (as time goes by) does not cause forgetting (Keppel & Underwood, 1962; Waugh & Norman, 1965); rather, other mechanisms that occur over time lead to memory loss, such as resolving interference caused by irrelevant events (Brown, Neath, & Chater, 2007; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012).

The interference hypothesis posits that memory traces suffer from degradation whenever information stored in memory competes with other activated items. Interference can occur during encoding or retrieval stages in memory, and perhaps also during storage activities. The amount of interference that will cause memory loss is defined as a function of the number, similarity, and strength of other items competing in memory. More specifically, interference can be caused by previously stored items that disturb both the encoding and the retrieval of new items (i.e. proactive interference). The reverse phenomenon (i.e. retroactive interference) is also possible, when new items exert interference on already stored information (Jonides et al., 2008). Proactive and retroactive interference can be caused by similarity-based interference, in which the strength of memory traces is disrupted by the similarity to other items. Similarity between two items can also occur on the time dimension, that is, proximal items can interfere when they appear close in time. This type of interference is related to temporal distinctiveness, where the strength of the memory traces are determined by the ratio of their temporal distances from the time of retrieval (Brown et al., 2007). Finally, in novelty-based interference new information that is similar to the previously encoded information is prone to less interference as it is encoded with less strength (Kelley, Neath, & Surprenant, 2013).

The cause of forgetting is probably one of the most debated topics within the working memory domain, with little chance for this long-standing contentious issue to be overcome any time soon by contrasting verbal theories. However, modeling may progressively circumvent obstacles that have prevented this field from converging for more than half a century. Our intention in the present proposal is not to discuss the interference-decay debate, but to focus on one specific model in order to improve its features, namely the Time-Based Ressource-Sharing model (TBRS) of Barrouillet et al. (2004) which assumes that temporal decay plays an essential role in forgetting. Although, the interference-decay debate is crucial to the advancement of our understanding, and particularly so because the tenets of the interference account suggest that neither rehearsal nor a short-term memory component are necessary to account for forgetting (Lewandowsky, Duncan, & Brown, 2004; Oberauer & Lewandowsky, 2014; Surprenant & Neath, 2009), the debate is probably worth a dozen PhD dissertations. Here, we assume the strategy of conducting research based on the TBRS model by simply arguing that it is the first decay-based model which is able to offer precise quantitative predictions, whereas previous theories based on decay could only observe the effect of time.

In the following sections, we will review a few models of working memory incorporating decay that can be considered as landmarks within the field. However, for the sake of brevity we have omitted several influential models in order to focus in greater depth on the TBRS model (Barrouillet et al., 2004), which is central to the present dissertation.

1.2.1 Early research on working memory: the modal model

Human memory is generally described by cognitive psychologists as a set of different components through which information is transferred. For instance, the conception of short-term and long-term memory as two distinctive stores (e.g. Atkinson & Shiffrin, 1968; Miller, 1956) has been the foundation of memory research. Based on this dichotomy, the modal model (also known as the multi-memory model and the multi-store model) of Atkinson and Shiffrin (1968) conceived selective attention as the main process allowing sensory information to be encoded temporarily before it is passed on to a more permanent memory store. Short-term memory in this model was described as a transient and passive function to store newly acquired information.

More specifically, Atkinson and Shiffrin's modal model (1968) assumes that human memory comprises a sensory register, a short-term memory store, and a long-term memory store. Sensory information (e.g. visual or auditive) coming from the environment is maintained only for a few hundred milliseconds in the sensory register, and unless this input is transferred to the next level, it rapidly decays and becomes inaccessible. Temporal constraints also apply to the second level of the model, as information stored within the short-term store suffers from rapid decay (within 15 to 30 seconds) if not actively maintained by control processes (a phenomenon that will be described in more detail later on). The short-term store is fueled by information coming from either the sensory register or the long-term memory store. The final level of this model, the long-term memory store, is presumed to be more or less limitless in its capacity and duration of retention, although estimates of these limits can only be roughly estimated (e.g. dozens of thousands of words, years). The transfer of information from the short-term memory to the long-term memory store is in part regulated by control processes, such as rehearsal. However, the authors note that information may also be transferred automatically. To support this claim, Aktinson and Shiffrin cite the work of Hebb (1961) and Melton (1963) that showed implicit learning from participants even when they are not actively trying to memorize information for later use.

According to the modal model, different control processes are associated to each component. The use of a specific control process is also defined in part by the ongoing task, its instructions, and on individual factors (e.g. experience of the participant). The authors note that the variety of control processes is unlimited since they are considered as an ephemeral phenomenon under a participant's control. In other words, they may include any coding technique or mnemonic strategy used to enhance the memorization of information or its retrieval. For instance, the function of controlled processes within the sensory register involves the decision making about which environmental outputs are selected and transferred to the short-term memory store.

Control processes associated with short-term memory include among others, search, storage, and retrieval strategies. The search process allows to locate a particular item within the short-term store in a recall task. Searching through the items stored in short-term memory is supposedly a rapid process (around 40 ms per item) because memory traces fade quickly with the passage of time. Another control process central to storage activities in short-term memory is rehearsal. This control process enables to maintain a limited amount of information in short-term memory in an immediate state of activation for a longer delay. By doing so, rehearsal also incidentally strengthens the memory trace by allowing other storage and coding processes to have extra time to operate. Another control process allowing to strengthen the memory trace is coding. The authors define this control process as "(...) a select alteration and/or addition to the information in the short-term store as the result of a search of the long-term store. This change may take a number of forms, often using preexisting associations already in long-term memory store." (Atkinson & Shiffrin, 1968, p. 39). Thus, short-term memory and long-term memory entertain close connections in this model, since according to the authors, information in the short-term store comes directly from the long-term store and indirectly from the sensory register. For example, if the word cat is visually presented, it may be coded as an auditory-verbal unit to strengthen the memory trace. However, for this to happen, a long-term memory search must take place to match the visual representation with the visual image of the cat.

The control processes associated to long-term memory participate either to the transfer of information from the short-term to the long-term store (as seen in our previous example), or they enable the search and retrieval of information from long-term memory. According to the authors, because the amount of information is important in long-term memory, the search process is more elaborate in this store compared to the short-term memory store. There can be indeed high latency times associated to a search in long-term memory when the information sought is temporarily unavailable until eventually the correct answer is found.

In the later version of the modal model (Atkinson & Shiffrin, 1971), short-term memory took a more predominant role as it was now considered as a temporary working memory, akin to consciousness, where information could be maintained with the help of controlled processes that are gathered in the short-term memory store, as seen in Figure 1.3. In that sense, the authors considered in this latter version of the model that short-term memory is an activation of some parts of long-term memory. To this day, this unitary view[3] of short-term and long-term memory is considered a serious option in several contemporary models of working memory (e.g. Cowan, 1999; Engle, Tuholski, Laughlin, & Conway, 1999).

Although the modal model seemed relevant at the time, it failed to account for several phenomena that outdated it rapidly (Baddeley & Hitch, 2007). Evidence against the idea of information simply passing through a passive store in short-term memory was provided by Craik and Lockhart (1972). In their article, the authors pointed out that a key factor for the creation of long-term memory traces is how much the newly acquired information has benefited from processing. For instance, words are better retained when participants categorize them actively instead of simply reading them in a list, as categorization makes individuals more active in the task at hand (Tekin & Roediger III, 2020). At the same time, long-term

[3]The conception of short-term (or working memory) and long-term memory as a continuum rather than structurally separate entities is still to this day a matter of debate (see Surprenant & Neath, 2009). However, we will not discuss this issue further as it is not essential to the rest of this thesis.

memory research took another turn with the distinction between semantic and episodic memory suggested by Tulving et al. (1972), and the importance of the level-of-processing effect became even more clear (Craik & Tulving, 1975; Tulving & Thomson, 1973).

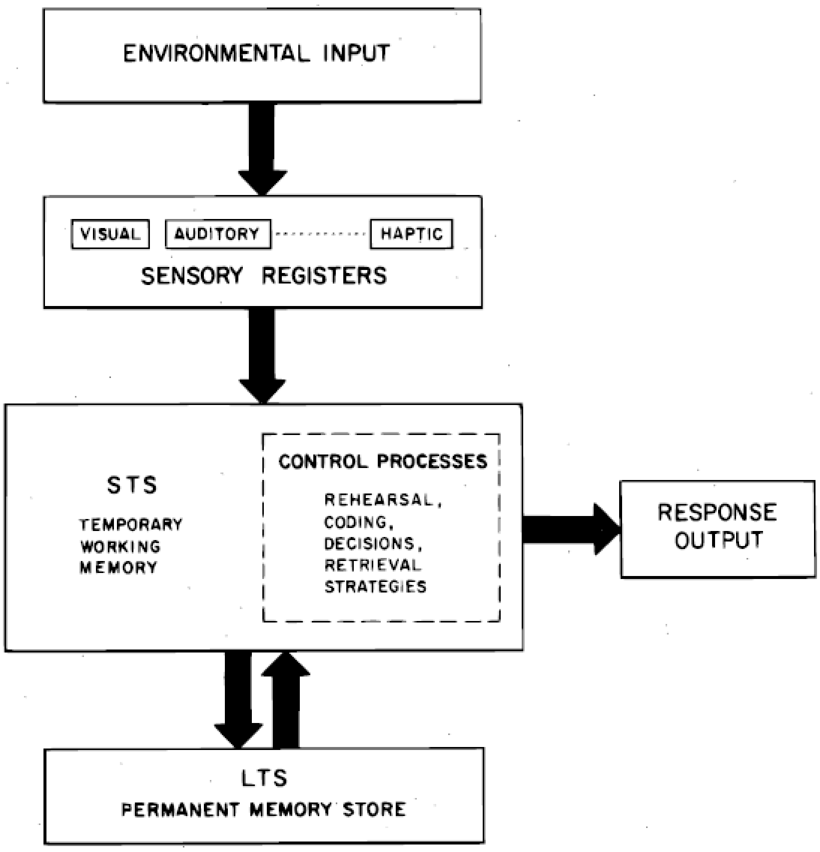


Figure 1.3: The structure of human memory according to the modal model from "The control processes of short-term memory", by R. C. Atkinson R. M. Shiffrin, 1971. Reprinted with permission.

Other controversies pleading against the modal model emerged from the field of neuropsychology in the same decade. The infamous case of the patient K.F. presenting a massive short-term memory impairment but with a functional long-term memory (Shallice & Warrington, 1970; Warrington & Shallice, 1969) was a living proof that the modal model was not adequate. Moreover, the fact that K.F. had preserved learning, memory, and comprehension abilities despite this short-term memory impairment was interpreted as contradictory evidence for a unitary short-term memory mechanism supposedly handling both storage and processing functions as conceptualized by the modal model.

In order to investigate further whether short-term storage and general cognitive activities are in fact managed by a unique system, Baddeley and Hitch (1974) ran a series of experiments using a dual task methodology where participants were asked to perform a second task (e.g. reasoning) while concurrently maintaining a memory load (i.e. a set of digits or letters). Although the results of these experiments showed an overall detrimental effect of memory load on the processing tasks, the far from dramatic decrease was interpreted as suggesting only a partial overlap between the systems responsible for maintenance and processing activities. This led Baddeley and Hitch (1974) to remodel the unitary system proper to the modal model into an active multicomponent system enabling storage and processing.

The modal model in summary

- Inspired by previous models emphasizing the dichotomy of short-term and long-term memory, the modal model of Atkinson and Shiffrin (1968) conceives human memory comprising a sensory register, and distinct short-term memory and long-term memory stores.
- Later, Atkinson and Shiffrin (1971) revisited the separation between short-term and long-term memory, and proposed a unitary conception of memory. This view is shared by several contemporary models.
- The modal model (Atkinson & Shiffrin, 1971) emphasized the role of the short-term store, described as a temporary working memory, akin to consciousness.
- In the modal model (Atkinson & Shiffrin, 1968, 1971), control processes play an important role in the maintenance of information. These control processes include among others search, storage, and retrieval strategies.
- Several controversies emerging especially from the field of neuropsychology plead against the modal model.
- The desire to tackle the controversies and to test the hypothesis of a unitary short-term memory mechanism supposedly handling both storage and processing functions as conceptualized by the modal model was pursued by Baddeley and Hitch (1974).
- The work of Baddeley and Hitch (1974) was a turning point in memory research that induced a conceptual transition from a passive short-term memory to an active working memory.

1.2.2 A turning point for working memory research: the Multicomponent Model

According to Baddeley and Logie (1999), the complex functioning of working memory enables us to apprehend our environment in the present moment. Based on information gathered from external stimuli and internally generated information, it enables us to perform complex cognitive activities such as solving problems by manipulating information on the spot, planning and executing actions. Moreover, working memory is considered the most direct hub to long-term memory and as such it could have a primary role in the consolidation of knowledge.

In their seminal work, Baddeley and Hitch (1974) conceptualize working memory as a multicomponent system that includes an attention controller and two subcomponents specialized in the temporary storage of material within a specific domain (the loop and the sketchpad). Over the years the model evolved, and a fourth component called the episodic buffer was eventually added (Baddeley, 2000). Figure 1.4 illustrates these two versions of the model.

The first domain-specific subcomponent, the phonological loop, initially called the articulatory loop, includes a temporary store and a (sub)vocal rehearsal system that enables the maintenance of information through repetition. Thus, visually presented items can also be stored within the phonological loop by the process of articulatory naming. Evidence in favor of the existence of a temporary storage and rehearsal system was demonstrated by the phonological similarity and the word length effects. Discovered by Conrad (1964), the phonological similarity effect refers to the phenomena where phonologically similar items (e.g. letters, words) are less accurately recalled in immediate serial recall tasks compared to dissimilar lists of items (e.g. *c, b, d,* and *v* versus *c, r, m* and *k*). This effect was interpreted as a manifestation of the the phonological nature of the short-term store. As to the articulatory rehearsal system, evidence was brought by the word length effect, which refers to the finding that recall accuracy declines as a function of the number of syllables that a word contains in immediate serial recall tasks. For instance, the serial recall of the words *association, opportunity, representative, organization, considerable, immediately, university,* and *individual* is less accurate than the recall of the following items: *sum, hate, harm, wit, bond, yield, worst,* and *twice* (Exp. 1, Baddeley et al., 1975). This finding proved to be robust even when lexical, semantic and frequency characteristics of the to-be-remembered words

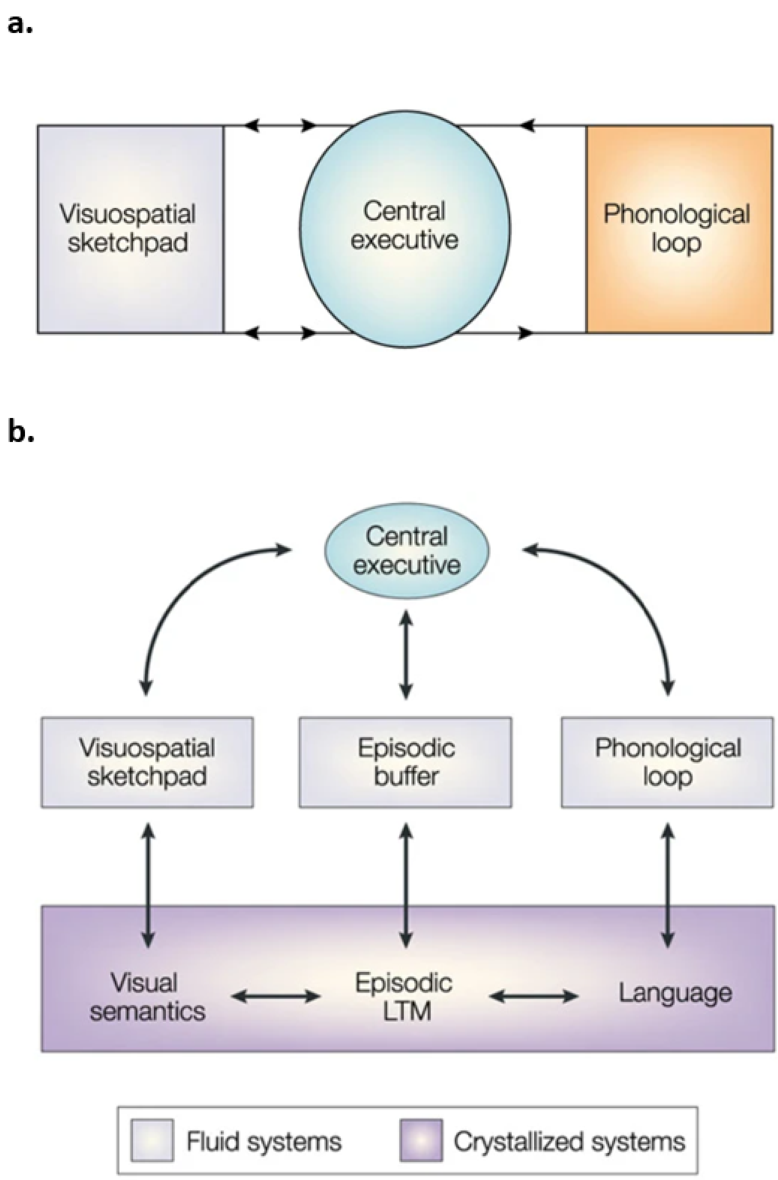


Figure 1.4: The Multicomponent model of working memory structure. Figure a. represents the original model as conceived by Baddeley and Hitch (1974). Figure b. illustrates a further development of the model, in which the episodic buffer has been added (Baddeley, 2000). The dark purple areas illustrate long-term or crystallized knowledge. Both figures reprinted from "Working memory: looking back and looking forward", by A.D. Baddeley, 2003, Nature Revue Neuroscience, 4, p. 830 and 835 (https://doi.org/10.1038/nrn1201). Copyright © 2003 by Nature Publishing Group.

were controlled for. Furthermore, within disyllabic words, items containing long vowels were less accurately recalled compared to words containing short vowels (e.g. *coerce, humane* versus *wicket, wiggle* - Exp. 3, Baddeley et al., 1975). Since rehearsal appeared to prevent the rapid loss of information (otherwise occurring after about only two seconds), the authors concluded that the linear relation between word length, articulation speed, and working memory capacity seems to support trace decay rather than interference as the source of forgetting.

Evidence from Cowan et al. (1992) work demonstrated, however, that the word length effect is not only due to the fact that long words take more time to rehearse but also that the time to articulate them is longer during recall. Further analysis of serial position revealed indeed that the performance of forward and backward recall of the first and last items to remember were better than what would be expected on the grounds of decay theory. Although the results were overall compatible with the phonological loop model of Baddeley (1986), the authors pointed out that decay theory may not explain all the variance in serial recall tasks, and most likely item distinctiveness could also play a role in recall performance (Murdock, 1960; see also Brown et al., 2007 for a model of distinctiveness). Around the same time, the phonological loop faced other challenges such as the fact that the model could not address how serial order is processed (see Henson, 1998), neither its relation to long-term learning, even though further research had shown that the loop might have an essential role in language learning (Baddeley, Gathercole, & Papagno, 1998). These issues were tackled by N. Burgess and Hitch (1992, 1999, 2006) with a computational approach applied to the multicomponent framework. Since their initial network model, several other competing models have been developed (e.g. Brown, Preece, & Hulme, 2000; Henson, 1998; Page & Norris, 1998) and short-term and long-term learning modeling has proven to be a fruitful area of research and remaining controversies.

Nonetheless, evidence in favor of the phonological loop model was also provided by the articulatory suppression effect. Participants that repeatedly recite an irrelevant word such as *the* during the presentation of the to-be-recalled items show a reduced working memory performance in contrast to the condition where no articulatory suppression is required from the participant. When the memoranda are visually presented the phonological similarity effect disappears (Murray, 1968), indicating that the articulatory suppression blocks rehearsal and item recoding. For auditorily presented stimuli, the phonological similarity effect remains (Baddeley, Lewis, & Vallar, 1984). Regarding the word-length effect, articulatory suppression appears to remove the phonological similarity effect, whether the items are presented visually or auditorily. This finding suggests that rehearsal is indeed responsible for the effect (Baddeley et al., 1975). Some controversies about the phonological loop were brought by the irrelevant sound effect. Participants that are exposed to irrelevant sound during an immediate serial recall task show impaired recall of visually presented information, even though they are explicitly instructed to not pay attention to the irrelevant sounds (Colle & Welsh, 1976). According to Baddeley (1986), this phenomena is the result of interference from the irrelevant sounds presented and the items to be remembered in the phonological loop. However, some studies on the influence of the irrelevant sound effect have appeared to contradict the multicomponent model as they found that the effect in question actually eliminates the phonological similarity effect (e.g. Estes, 1973; Murray, 1968). On the other hand, other studies have indicated robust phonological similarity effects even though participants are exposed to irrelevant sounds (e.g. Colle & Welsh, 1976; Salame & Baddeley, 1982), provided that the number of items to be maintained in short-term memory do not exceed the capacity of the phonological loop (Salamé & Baddeley, 1986). This hypothesis of abandonment of the phonological loop by participants in longer list lengths is, however, contested by some authors (e.g. Jones, 1993; Jones, Hughes, & Macken, 2007). Other alternative explanations emerged, among others the changing state hypothesis of Jones (1993), according to which only changing sounds will impair recall.

The second domain-specific subcomponent, the sketchpad, initially named the scratchpad, is dedicated to the storage of visuospatial information. In the early version of the multicomponent model, the sketchpad was thought to store only spatial information, but the work of Logie (1986) showed that the subcomponent could deal with visual information of all sorts. Evidence points out to the existence of several components within the visual sketchpad. One of these components, the visual cache, allows the retention of visual patterns, while the other one, known as the inner scribe, enables the memorization of sequence of movements (Baddeley & Logie, 1999). The study of Logie and Marchetti (1991) is a nice demonstration of a double dissociation between these two subcomponents. In this work, the authors showed that the retention of spatial patterns can be hindered by arm movements executed between retention intervals, whereas this action did not have any effect on the visual information to be memorized. In contrast, a visual interference task presented between the display of items to remember and their recall showed a disruption of the maintenance of the visual information while the maintenance of the spatial patterns was not affected by this manipulation.

The attention controller known as the central executive, has undergone several revisions since its initial conception. The original conception of the construct was indeed so broad that it was considered as a homunculus (Baddeley & Logie, 1999). Inspired by the field of attention[4], the central executive was defined as an attention-based component enabling several functions: attentional focus, divided attention and task switching, as suggested by Baddeley (1996) and by Baddeley and Logie (1999). The fourth function of the central executive was thought to make working memory and long-term memory work together. However, this function was later on alienated (Baddeley, 2000), and it is now integrated as a process in the episodic buffer.

Following this seminal work, ample evidence has been discovered in favor of the involvement of attentional focus in working memory (e.g. Allen, Baddeley, & Hitch, 2014; Morris & Jones, 1990). For instance, (Robbins et al., 1996) showed in a series of experiments that articulatory suppression has no effect on the recall of briefly presented chess positions in players whereas a concurrent random letter generation task that requires attentional resources impaired working memory performance.

Evidence for the capacity to divide attention was also investigated through neuropsychological studies involving clinical populations (Baddeley, Della Sala, Papagno, & Spinnler, 1997; Baddeley, Logie, Bressi, Della Sala, & Spinnler, 1986). For example, in the study by Baddeley et al. (1986) participants were required to memorize strings of random digits while tracking a moving target on a computer screen. In order to evaluate specifically the cost of divided attention, each task was first performed separately and then concomitantly. Results showed that in healthy adults, performance in dual task decreased by 20% to 25% compared to the single-task performance in both tasks. In contrast, patients suffering from Alzheimer's disease showed a drop down of 60% in their performance when performing both task at the same time compared to their performance in a single task, indicating thus a specific impairment of divided attention in these patients.

Regarding the capacity of switching, Baddeley, Chincotta, and Adlam (2001) investigated the role of the central executive in switching attention between two tasks through a series of experiments involving a dual-task methodology. The results

[4]The central executive was inspired by the field of attention, and specifically by the model of Norman and Shallice (1986) that suggested two distinctive mechanisms of attention allowing both automatic actions and controlled actions. The construct of the central executive was influenced in particular by the Supervisory Attentional System thought to intervene when the automatic system fails.

showed that in a single task, the phonological loop plays an important function in switching and even more so than the executive. According to Baddeley and Hitch (2007), switching may occur in several ways, and executive resources are not necessarily involved in all of them.

Finally, the youngest addition in the multicomponent model, the episodic buffer, is defined as an interface between the episodic long-term memory and working memory. The episodic buffer is an attention-based limited storage system for information coming from different sources (i.e. visual, verbal) enabling a multi-dimensional encoding. The buffer is thought to interact with the different components of the model. For instance, long-term knowledge such as frequency, semantic, characteristics of words such as concreteness will affect how the items are recalled in working memory (e.g. Allen & Hulme, 2006; Gathercole & Adams, 1994). For instance, language habits will influence chunking abilities, allowing thus to increase significantly the span if the participants are able to extricate meaning from the memoranda (Baddeley, Hitch, & Allen, 2009). Similarly, verbal information forming a meaningful sentence or displayed in familiar visuospatial configuration are known to enhance working memory capacity (e.g. Baddeley et al., 2009; Darling, Parker, Goodall, Havelka, & Allen, 2014). In sum, Baddeley, Hitch, and Allen (2020, p. 28) acknowledge the complex interactions between working memory and long-term memory, however they note that "(...) the assertion that working memory is simply activated LTM (long-term memory) is not erroneous, but is unhelpful, other than a placeholder for further research."

Overall, the multicomponent model of working memory established a structure of general functions providing interpretations of experimental data. Their work profoundly changed the theoretical landscape of the field and lead other theorists to explore further the interplay between storage and processing in working memory.

The multicomponent model in summary

- The multicomponent model of Baddeley and Hitch (1974) conceptualized working memory as a system that comprises a central executive and two subcomponents specialized in the temporary storage of material within a specific domain (the loop and the sketchpad). Later a fourth component called the episodic buffer was added (Baddeley, 2000).

- The phonological loop includes a temporary store and a (sub)vocal rehearsal system that enables the maintenance of information through repetition. Evidence in favor of the loop was brought by the phonological similarity effect, the word length effect, the articulatory suppression effect and the irrelevant sound effect.
- The sketchpad is dedicated to the storage of visuospatial information. Within the sketchpad, the visual cache allows the retention of visual patterns, while the inner scribe enables the memorization of sequence of movements.
- The central executive has undergone several revisions since its initial conception. It enables several functions: attentional focus, divided attention and task switching (Baddeley, 1996; Baddeley & Logie, 1999).
- The episodic buffer is defined as an interface between the episodic long-term memory and working memory, that allows limited storage for information coming from different sources.
- The multicomponent model transformed the theoretical landscape and led other models to explore further the interplay between storage and processing activities.

1.2.3 Theoretical aspects of the TBRS from developmental studies

Following Baddeley and Hitch's work (1974), the view of separate systems responsible for storage and processing functions was less favored in subsequent models. Several empirical findings were interpreted as consistent with a unitary view of short-term memory where processing and storage functions are fueled by a shared common resource that causes a trade-off between these functions (e.g. Anderson, Reder, & Lebiere, 1996; Barrouillet et al., 2004, 2007).

In line with this trade-off conception, Case, Kurland, and Goldberg (1982) argued for a shared cognitive space between storage and processing activities within short-term memory. According to the authors, the complexity of a given task determines how much space it consumes. A processing task that is difficult to perform will monopolize much of this common space, resulting in storage activities being inevitably compromised. Evidence in favor of this conception was reported by Case et al. (1982) in a series of experiments in which performance of children in a counting span task were equated to the performance of adults who were required to perform the task in an unfamiliar language. In this counting span task, participants were asked to count dots in successive arrays and memorize the successive total number of dots for further recall. The conclusion of this work was that processing becomes more efficient with age, thus requiring less cognitive space, which in turn allows more space for storage activities. According to the authors, this additional available space for storage with age explains the developmental increase of memory span. It is worth noting in Case's et al. work (1982) that although processing efficiency is operationalized through processing speed, time does not play any other role than a mere indicator for efficiency. Inefficient processing (i.e. slow processing) reveals that attentional resources are exhausted by the processing task, leaving thus an insufficient amount of attention to storage.

Case's account was, however, challenged by Towse and Hitch (1995) who proposed an alternative narrative to explain the interplay between storage and processing functions within working memory. According to the authors, the developmental increase found by Case et al. (1982) was not due to a more efficient distribution of the available cognitive space, but rather it could be attributed to a faster processing ability with age that subsequently decreases the retention periods. Compared to adults, children are less expert in counting, thus they will take more time to finish the processing activity and as a result their memory traces will suffer from time related decay for a longer period of time. In sum, the authors duly noted that processing efficiency and retention intervals of the memoranda are confounded in Case et al. (1982) experiments.

According to Towse and Hitch (1995), resource-sharing and the notion of cognitive space that derives from it are unnecessary, as the sole hypothesis of temporal decay can explain the findings. To test this hypothesis, the authors designed a counting span task in which the cognitive cost of the processing task was manipulated. More specifically, the difficulty of the task varied as a function of the discriminability of the processing items. Simply put, they implemented an easy condition involving processing targets that were recognized and processed faster than in a difficult condition where distractor items shared similar visual features with the processing targets. Additionally, a third condition labeled combined was designed by using the processing items of the easy condition, but with an increased amount of items to process in order to equate the duration of the trials to the difficult condition. This allowed the authors to disentangle the processing speed of the task and the duration of the retention periods from each other. The results from participants aged from 6 to 11 years old showed that recall performance was higher for the easy condition compared to the difficult and the combined condition. Also, recall performance in the difficult condition and the combined condition did not differ from each other. This pattern of results was constant across all ages. These findings were in line with the predictions of Towse and Hitch (1995) that posited that recall performance is constrained by temporal factors only and not by the difficulty of the task. Finally, to account for these findings, the authors suggested a task-switching hypothesis, according to which participants alternate during a trial between storage and processing activities.

Further studies conducted with a variety of working memory tasks (e.g. operation and reading span task) have found results that concur with the task-switching hypothesis, viz. recall performance is a function of the retention periods, with longer retention intervals resulting in poorer working memory spans (Towse, Hitch, & Hutton, 1998, 2002). The mechanisms allowing to maintain the memoranda was, however, unspecified. Nevertheless, the authors ruled out one possible explanation: the phonological loop. It was clear that a rehearsal mechanism enabling to hold information for a couple of seconds could not explain why the participants were unable to consolidate items with a greater amount of available information.

Within this theoretical context, Barrouillet and Camos (2001) explored further the two alternative accounts (i.e. Case's cognitive space hypothesis *vs.* Towse and Hitch's memory decay hypothesis) with a series of experiments. Essentially, Barrouillet & Camos (2001, p.4) argued that the interplay between storage and processing as proposed by Towse et al. (1998) was an oversimplification, since " (...) it is possible that the counting of larger arrays results not only in longer times, but also in a higher cognitive load." The experimental design involved tasks in which the processing events were identical in terms of duration but differed in terms of the cognitive cost associated to it. In the first experiments, children aged of 8 and 11 years old carried out a counting span task where they had to count red target dots mixed with green decoy dots. After each processing episode a letter was presented. At the end of the trial, the participants were asked to recall the letters in their correct order of presentation. The list length of the storage-processing events varied from 2 to 6. The counting time of each child was measured, which allowed to tailor an individualized "baba span task". The baba span task was completed by the same children a few weeks later. This task was essentially identical to the aforementioned counting span, except that the children were being asked to say repeatedly “baba" when an empty screen was presented instead of the counting task. As the baba span task requires the automatic repetition of the same syllable, it was considered that a lower cognitive cost was associated to it, compared to the counting span. Thus, this manipulation allowed the authors to have identical intervals of distractions in both tasks, while varying the cognitive cost associated to the processing tasks.

According to Case’s cognitive space hypothesis, the counting span should produce better recall performance as it is a more demanding task compared to the baba span task. In contrast, Towse and Hitch’s memory decay hypothesis predicts no difference between the two tasks, as the retention intervals are identical in both tasks. Barrouillet and Camos (2001) reported findings favoring the task-switching model, as no difference between the counting span and the baba span task were found. Further experiments involving roughly the same experimental design with younger children (considered as less expert in counting) showed similar results.

Unconvinced by these results that seemed to indicate that temporal decay alone might explain the ins and outs of working memory, Barrouillet and Camos (2001) tested in a third experiment again the same hypothesis but this time the cognitive cost associated to the processing task was dramatically increased. Instead of counting, the children were asked to complete either an operation span task or a baba span task. The operation span task consisted of solving and verifying the result of additions (e.g. $8 + 7 = 15$? ; $4 + 7 + 8 = 19$?) presented on the screen. The participants were required to respond accordingly either “true” or “false”, which triggered the presentation of the next storage item. In sum, in the condition involving a high cognitive cost, a series of consonants were interspersed with mathematical operation, whereas in the low cognitive cost condition, the operations were substituted with empty screens that cued the participant to repeat baba. The duration of the baba repetitions corresponded to the mean processing time of the group performing the operation span task. This time the results favored Case’s hypothesis, as they showed an increased performance in the baba span task compared to the operation span task, suggesting a trade-off between storage and processing.

In light of these results, Barrouillet & Camos (2001, pp. 15-16) emphasized that “(...) a more sophisticated model is needed because it seems that two kinds of limitations have to be taken into account in working memory functioning. The first is a limitation in the time a given piece of knowledge can be kept active in working memory, as Towse, Hitch, and Hutton (1998) have demonstrated. The second limitation could result from a limited pool of resource, whatever we mean by 'resource' (capacity for activation, for controlled attention, etc.), as suggested by the difference between operation span and baba span even when time parameters are held constant."

It is worth noting that the finding of Barrouillet and Camos (2001) indicated only a *moderate* effect of the cognitive cost of the task on recall performance. Therefore, how the cognitive system was able to maintain the memoranda despite a demanding concurrent processing activity still remained an open question. Covert retrieval, defined as a rapid memory search process that allows the reactivation of deactivated items (e.g. Cowan et al., 1994), was considered as an interesting candidate. According to Barrouillet and Camos (2001) the reactivation of memory traces might explain how participants are able to recall the memoranda and counteract decay even when the use of rehearsal strategies is prevented by the repetition of the baba syllable or by counting. In addition to the phenomenon of covert retrieval, Barrouillet and Camos (2001) also suggested an additional hypothesis to explain the surprisingly slow rate of decay of the memory traces: participants alternate between storage and processing activities much more often than what was originally conceived by Towse et al. (1998). If this were the case, then switching would not only be influenced from the inherent schedule of the complex span task (i.e. participant switch from processing to storage and vice versa only when cued to do so by the task), but also during the processing activity itself.

Overall, the preliminary studies of Barrouillet and Camos (2001) revealed a more complex reality than what was initially envisioned by either the tenants of the resource-sharing or decay-based models. The findings that we described above led the authors to suggest that working memory functioning is not only constrained by a limited attentional resource but also by temporal factors. These premises are the core foundation of the time-based resource-sharing model of Barrouillet et al. (2004) that we will describe in the next section.

1.2.4 Time-Based Resource-Sharing model

As discussed above, the notion of resource-sharing is not new, but the TBRS model argues for a more elaborate account of the trade-off between storage and processing functions while taking into consideration temporal factors (Barrouillet & Camos, 2014b).

The TBRS model is based on four main propositions. First, the model assumes that the two main functions of working memory, storage and processing, rely both on a single limited resource: controlled attention. While it is commonly recognized that all complex processing activities, such as reading, problem solving or computation, require controlled attention (e.g. Engle, 2002; Kane & Engle, 2003), Barrouillet et al. (2004) posit that simpler activities, such as reading aloud digits or letters, also require attention. Moreover, the TBRS assumes that maintaining information requires attentional resources as well, thus the proper functioning of working memory necessarily involves resource sharing.

According to the second hypothesis, when attention is not focused on maintenance activities, memory traces suffer from decay as time goes by. The memory traces can, however, be reconstructed by recruiting attention (Barrouillet & Camos, 2014b). This reactivation can occur via two distinctive maintenance mechanism. (1) A general maintenance process of rapid recovery through attentional focus, called refreshing, and (2) a specific maintenance mechanism based on articulatory rehearsal, referred to as the articulatory loop, where the articulatory mechanism is connected to a motor buffer that maintains the motor program devoted to speech production (Barrouillet & Camos, 2020) [5]. Refreshing and the articulatory loop contribute both to counter the spontaneous decay of verbal information (Camos, Lagner, & Barrouillet, 2009).

The third assumption of the TBRS model is that activities of storage and processing in working memory can only occur sequentially. When attention is devoted to processing, it is no longer available for maintenance, thus working memory traces suffer from time-based decay and representation-based interference. Conversely, when attention is occupied by maintenance activities, it cannot be available for

[5]The articulatory loop in the TBRS model has undergone a recent theoretical redefinition as it was initially conceptualized in Barrouillet and Camos (2014b) as a phonological loop initially suggested by Baddeley (1986) in which the articulatory mechanism is connected to a phonological buffer.

processing activities. According to Barrouillet and Camos (2014b) the well-known general hypothesis of a central bottleneck (Pashler, 1999) concerning a variety of cognitive processes is a valid candidate to explain the sequential nature of storage and processing activity. According to this view, when the bottleneck is busy, no further processing can be carried out in parallel. This third assumption could thus simply be a sequel to the psychological refractory period referring to the idea that attention can only handle pieces of information one at a time (Pashler, 1994). An alternative hypothesis to explain this phenomena, adopted by concurrent theories of working memory (e.g.Oberauer, 2002), is the notion of a limited focus of attention enabling the processing of only one piece of information at a time. Whether it is the bottleneck hypothesis or the limited capacity of the attentional focus that constrains working memory functioning, the core idea conveyed by the authors is that storage and processing cannot occur simultaneously. This means that when engaged in a processing activity, such as attempting to add up successive numbers, the maintenance of newly acquired information is necessarily postponed until the end of the processing activity. This brings us to the last assumption of the model.

On the grounds that attention is a serially shared resource, the final assumption of the model is that the diverse activities must alternate or switch between one another. However, unlike the task-switching model of Towse and Hitch (1995); Towse et al. (1998), this switching occurs not only when processing and storage events are displayed during a complex span task but also during a processing episode. For instance, during a processing task involving the resolution of mathematical operations, participants may rapidly switch back and forth between processing and storage activities when intermediate results are found. This idea of rapid switching between the two main function of working memory, without compromising neither the processing nor the storage task, led the authors to submit a new conception of the notion of cognitive load.

Cognitive load in the TBRS model. Introduced by Barrouillet et al. (2004), the cognitive load is the main derivation of the four above assumptions. The concept of cognitive load is present in several working memory models (e.g. Case et al., 1982) and classically refers to the complexity of the task to be performed. In other words, the more complex a secondary task is, the more cognitive resource its execution will require, leaving thus only little resources available to maintain the memory traces. This is at odds with the conception of cognitive load in the TBRS model, in which the complexity of a task is irrelevant. In fact, the complexity of a task can be included in a more general cognitive load factor that assumes that the proportion of time during which a given task occupies attention prevents maintenance of memory traces. This novelty is supposed to solve the confound that a more complex task takes longer to complete, and it is the most direct assumption of the decay hypothesis.

Simply put, the Cognitive Load (CL) can be defined as:

$$CL = Nt/T \tag{1.1}$$

N = Number of items to process after each storage item
t = Total time of attentional capture
T = Total time to perform the task

Barrouillet et al. (2004) found evidence for a linear relationship between cognitive load and working memory performance. Thus, the greater the proportion of time during which attention is fully occupied (i.e. the higher the cognitive load), the lower the memory performance is. Figure 1.5 illustrates this phenomenon. Numerous studies have demonstrated strong evidence in favor of the cognitive load effect (e.g. Barrouillet et al., 2004, 2007; Barrouillet & Camos, 2012; Barrouillet, Portrat, & Camos, 2011).

Experimentally the cognitive load can be manipulated in several ways. For instance, in one study conducted by Barrouillet et al. (2004, Exp.2), participants completed a reading digit span task consisting of either 6 or 10 digits to be read interspersed with consonants to be remembered. The authors varied the cognitive load by manipulating the number of items to be processed while keeping the inter-letter intervals constant. This resulted in a slow pace in which the processing items (i.e. digits) where presented at 1000 ms each, and a fast pace where these items were presented at only 600 ms per item. Further experiments involving processing tasks to be performed silently confirmed that this cognitive load effect is not explained simply by the fact that participants are reading digit aloud, and thus results cannot be explained in terms of the articulatory suppression effect (Lépine, Barrouillet, & Camos, 2005).

Other than increasing the amount of processing steps to complete within the same time frame, cognitive load can also be varied by increasing the duration of the attentional capture. For example, in a replication of Experiment 1 of Barrouillet et

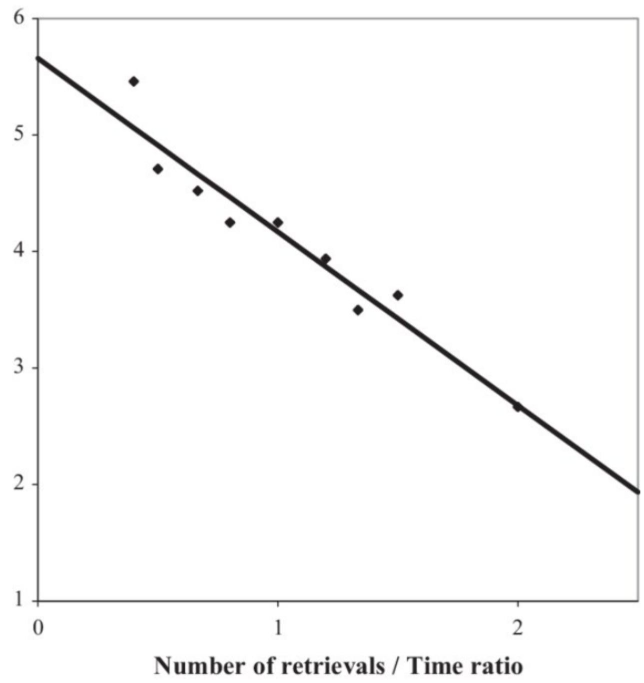


Figure 1.5: Mean working memory spans as a function of the cognitive load depicted in terms of the number of digits (or retrievals) to be read per seconds. Reprinted from “On the Law Relating Processing to Storage in Working Memory" (p. 179), P. Barrouillet, S. Portrat and V. Camos, 2011. Psychological Review. Copyright © 2011, American Psychological Association.

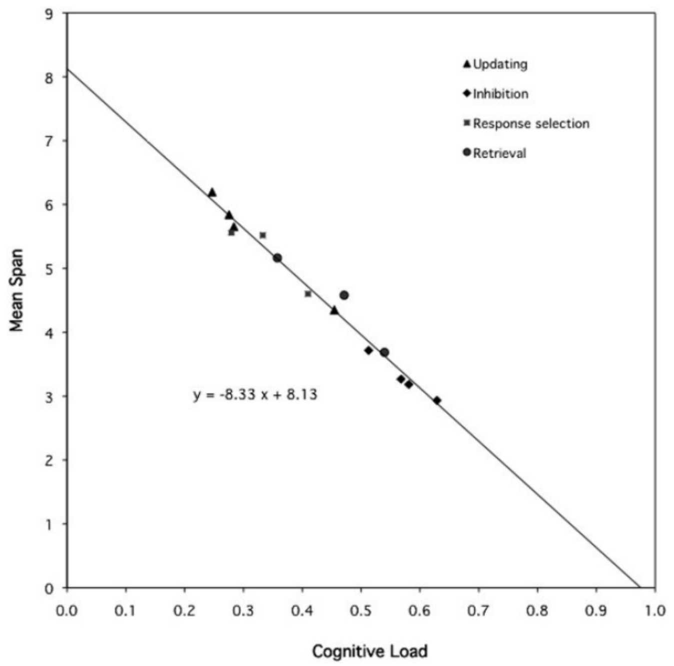


Figure 1.6: Mean span as a function of the cognitive load involved by four different types of tasks. Reprinted from “On the Law Relating Processing to Storage in Working Memory" (p. 182), by P. Barrouillet, S. Portrat and V. Camos, 2011. Psychological Review. Copyright © 2011, American Psychological Association.

al. (2007), the cognitive load manipulation was achieved by asking participants to read aloud digits either in their Arabic, Roman or written form, while keeping the time to complete the tasks constant. Since reading digits in their Roman or written form compared to their Arabic form is more time consuming, this created a high CL and low CL condition respectively, while keeping the processing tasks similar in terms of the cognitive processes that they require.

Finally, cognitive load can also be manipulated by simply varying the amount of time available to perform the task while keeping the number of processing steps and their attentional capture constant (e.g. Exp. 5-6, Barrouillet et al., 2004). Perhaps one of the most counter intuitive effect that derives from these manipulations is that a longer task to complete do not necessarily imply a higher cognitive load compared to a shorter task. Indeed, increasing the number of processing steps at a constant pace will not have any effect on the cognitive load of the task.

The cognitive load effect is a robust finding. It has been replicated numerous times in adults and children with or without school difficulties (Corbin, Moissenet, & Camos, 2012; Portrat, Camos, & Barrouillet, 2009). Moreover, the relation between cognitive load and performance remains constant across different attention-based processing activities if the ratio between attentional capture and the total duration to execute a task remains the same. For instance, processing activities as diverse as retrieval from long-term memory, updating, or inhibition in visuo-spatial or verbal tasks do not interact with the measure of CL in the TBRS, as seen in Figure 1.6. However, in line with previous findings that have reported a detrimental effect of choice reaction time (RT) tasks on serial recall performance, but not for simple RT tasks (Szmalec, Vandierendonck, & Kemps, 2005), processing tasks that require very little attentional resources will not have any effect on cognitive load. In other words, increasing the number of processing steps that require only stimulus detection do not influence the cognitive load of the task (Exp. 4, Barrouillet et al., 2007). This adds evidence to the idea that the raw duration of the task is not sufficient to explain these findings, what counts is the proportion of time during which attention is captured.

Cognitive load variations

There are three important factors in the CL equation: **N**, **t** and **T**. Thus, there are three different ways to vary the cognitive load in a task.

1. Vary the number of processing steps within the task
2. Vary the time to execute the task
3. Vary the attentional capture of the processing task

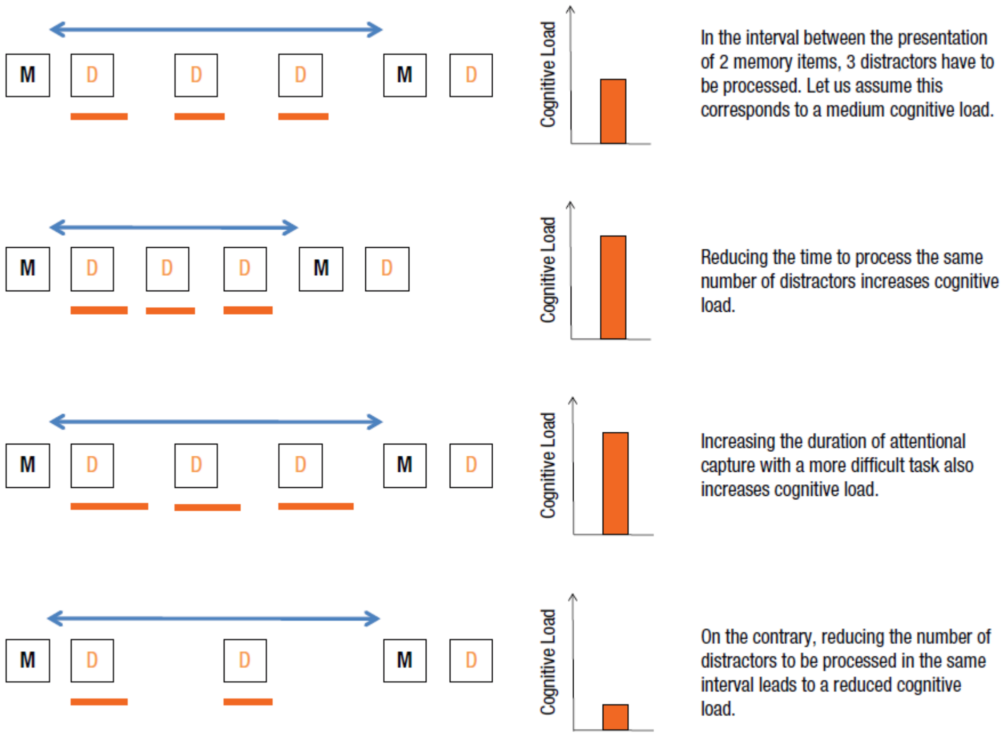


Figure 1.7: Illustration of different ways to manipulate the cognitive load of the processing task in a complex span task. The processing steps are represented by the boxes marked as *D*. Storage items are marked *M*. The horizontal orange bars illustrate hypothetical durations of attentional capture that results from each processing item, and the blue arrows illustrate the inter-letter intervals. The bar graphs represent the variations in cognitive load. Reprinted from "As Time Goes By: Temporal Constraints in Working Memory" (p. 415), P. Barrouillet and V. Camos, 2012. Current Directions in Psychological Science. Copyright © 2012, SAGE Publications.

TBRS in summary

- The TBRS model has four main assumptions:
  1. **Attention** is a shared resource between processing and maintenance activities in working memory.
  2. Memory traces spontaneously suffer from **decay** unless refreshed in working memory.
  3. All activities within working memory are **serial**.
  4. Alternation of activities in working memory involves **switching**.
- **Cognitive load**, the proportion of time during which a distracting task occupies attention, is the single factor that determines the span. The span is the lowest when the maintenance of memory traces is prevented by a high load of distractions.
- The TBRS model can account for a large range of working memory functions, but its core functioning still merits further attention, in particular to make more accurate predictions based on storage and processing alternation patterns.
- A to-be-tested hypothesis of the present dissertation is that storage and processing alternation patterns produce switch costs partially independent of cognitive load.

The architecture of the TBRS model. The previous section detailed the basic functional aspects of working memory as conceived by the TBRS model. From their empirical findings, the authors suggested a complete new cognitive architecture of working memory illustrated in Figure 1.8. The following section describes this architecture, which comprises of peripheral systems, an episodic buffer, and a production system. The latter ones, the episodic buffer and the production system, are thought to constantly interact with each other within a central system called the executive loop.

The peripheral systems contain several buffers that allow the maintenance of modality-specific information. Among these buffers, the authors posit the existence of phonological, visuo-spatial, and motor buffers. As already discussed in the previous section, visuo-spatial information does not profit from a domain-specific maintenance system. Such a system, however, exists for verbal information, as a limited amount of motor programs for articulation may be maintained via the articulatory loop. The peripheral system also includes modules of long-term memory that transfers knowledge for the construction of working memory representations. The peripheral system can also receive outputs from the central system, allowing for instance to store information through a learning process, or receiving the instructions to response productions enabled by the motor buffer. The authors do not exclude the possibility of additional buffers such as auditory, musical, kinaesthetic and haptic buffers, but they point out that the relation of these sensory inputs with working memory is less known, which explains why they are not explicitly included for the time being in the model. Overall, working memory representations can integrate information from the buffers and long-term memory. In that sense, all representations are episodic in nature, as working memory representations can incorporate elements of contextual knowledge that may change their meaning. The authors illustrate this point with a simple example on how the representation of a date on a calendar will be different than the one of a price tag even when they both involve the same number (e.g. 14th of December and 14 euros).

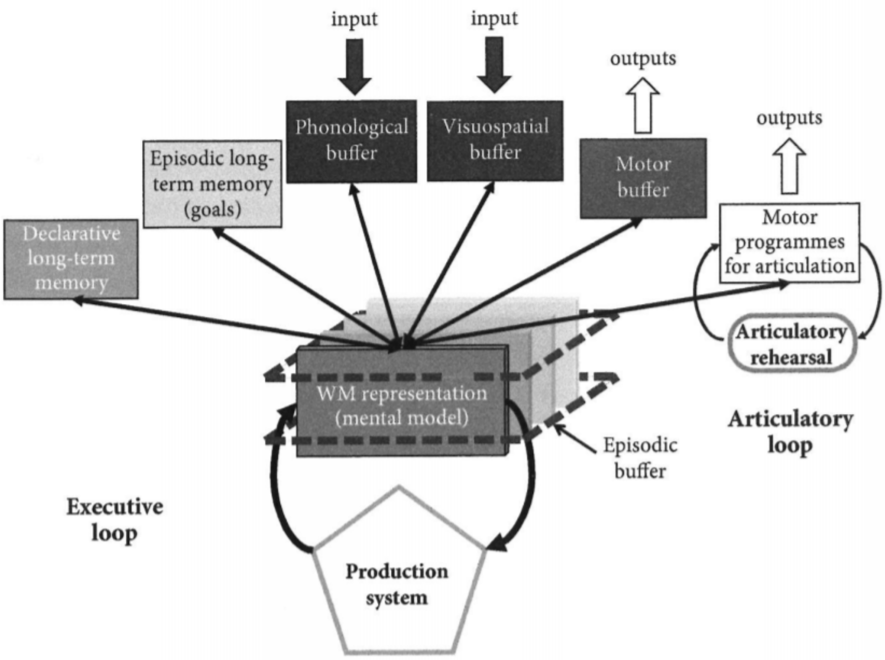


Figure 1.8: The cognitive architecture of the TBRS model. Reprinted from Working Memory: State of the Science (p. 101), by P. Barrouillet and V. Camos, 2020. Oxford University Press. Copyright © 2020 by Oxford University Press.

Inspired by Baddeley's work (Baddeley, 2000), the episodic buffer can hold working memory representations from multiple sources. Although this episodic buffer can maintain up to four representations in an active state, only one can remain accessible at a time for active treatment. This is line with Oberauer (2002)'s conception, for whom only one representation can be held active in the focus of attention, while the rest of the memory content in the episodic buffer still have the privilege to remain more accessible than other non-active representations.

Regarding forgetting, according to the TBRS model working memory representations are subject to deterioration caused by temporal decay and interference. Within the episodic buffer, working memory representations are constantly scanned, maintained, or changed by procedural knowledge, referred to as productions by the authors. This constant synergy between the episodic buffer and the procedural system results in executive loops. According to the authors, “this central system [the executive loop] in which goal-directed-cognition takes place can be conceived as the seat of thought as it is in charge of the construction, maintenance and processing of multimodal transient working memory representations” (Barrouillet & Camos, 2020, p. 100).

The production system is in charge of reading working memory representations and it can decide the current action to pursue as a function of intentions. It is described as a structure that will deploy control structures such as “*if* conditions *then* action script”. More specifically, when a representation is read by the production system, a system of pattern matching will occur. This pattern matching process will select the action that best matches the representation and the goal in mind. This matching process results in the selection (and then in the application) of the most appropriate action. In other words, productions can be thought as a system which triggers the executive functions that operate on working memory representations. According to the authors, “Productions can reconstruct partially degraded representations for maintenance purpose, modify representations for reaching goals, switch between representations held in the episodic buffer, or retrieve information from long-term memory and sensory buffers for constructing new representations” (Barrouillet & Camos, 2020, p. 102). Only one production can apply to one working memory representation at a time. This means that processing and maintenance activities in working memory always function in a sequential manner. Thus, the refreshing schedule is also presumed to occur in a cyclic ordered fashion, although alternative schedules are not ruled out in definitive (Lemaire, Pageot, Plancher, & Portrat, 2018).

In a previous version of the TBRS model (Barrouillet & Camos, 2014b), the authors stipulated the existence of a specific goal module as part of the peripheral system. However, the authors in their revised version Barrouillet and Camos (2020) specify the presence of a final goal and sub-goals. The final goal would correspond to the objective of a complex activity, while the sub-goal would refer to a cognitive step needed before reaching the final goal. For instance, when solving arithmetic calculations, the final goal is to find the solution, whereas one sub-goal could be to find an intermediate result which will enable us to solve the mathematical operation. According to the authors, the final goals are not necessarily maintained actively in working memory. However, goals trigger the activation of knowledge in long-term memory, which will build the contextual frame of the task. This context is essential to guide the cognitive activity on the right track. Furthermore, the context could act as a cue enabling individuals to retrieve goals from long-term memory, which in turn could help reconstruct or update the context if necessary. The sub-goals are in turn created by the production system (for instance, first initiate a movement of the arm, then open the hand in view of grasping an object). However, because the cognitive activity constantly moves on to the next sub-goal until the final goal is reached, the maintenance of these sub-goals is only temporary.

Overall, as seen from the previous paragraph, the TBRS model conceives working memory and executive functions as closely related to one another. According to the authors " (...) executive control is in the TBRS model an emergent proprety of cognitive functioning " (Barrouillet & Camos, 2020, p. 103).

Limits of the TBRS model. Although Barrouillet and Camos have amassed strong evidence in favor of their model, some findings can be considered as non-consistent within their theoretical framework. For instance, the fact that not all memoranda can be refreshed in a recognition paradigm suggests that the cognitive load might in some case be task or material dependent (Ricker & Cowan, 2010; Vergauwe, Camos, & Barrouillet, 2014). This is the case for storage items consisting of unfamiliar characters or fonts of different colors, which are inevitably lost over time even in the absence of distractors. This questions the universality of the cognitive load effect, as for some items recall performance seems to depend solely on time-based factors. According to Ricker and Cowan (2010) and Vergauwe et al. (2014), one explanation for these finding is that long-term knowledge of the memoranda is required for attention-based refreshing to operate.

More recently, Ricker and Vergauwe (2020) failed to find a cognitive load effect across a series of experiments in visual working memory. As expected from previous work, tasks that used low-level perceptual features as memoranda were not sensitive to cognitive load manipulations. But, more surprisingly, these findings extended to conceptual memory materials (i.e. canonical angles), although previous work had already demonstrated cognitive load effects in visuo-spatial working memory tasks (e.g. Barrouillet, De Paepe, & Langerock, 2012; Langerock, Vergauwe, & Barrouillet, 2014; Vergauwe, Dewaele, Langerock, & Barrouillet, 2012).

Ricker and Vergauwe (2020) identified several potential explanations for this discrepancy. First, the absence of cognitive load effect might be due to the processing material and the tasks themselves. As stated above, visuo-spatial tasks have been used in previous work, however, this is the first study that has used an angle reproduction task. Moreover, the authors did not use a complex span task unlike the vast majority of the existing work on this topic. Instead, they used a Brown-Peterson task in which participants are first asked to memorize the items that are displayed successively one at-a-time, and the processing task is not executed before the presentation of all storage items. Recall that in complex span tasks, the presentation of storage and processing episodes is always interspersed. Thus, the crucial difference between the Brown-Peterson task and the complex span task is whether the participant is confronted to a unique or multiple retention intervals after the presentation of each storage item. However, according to the authors there is no theoretical reason for this number of retention intervals to matter.

Another explanation suggested by Ricker and Vergauwe (2020) for the unexpected failure to find a cognitive load effect in their study is related to the presentation time of the memoranda that may lead to consolidation of memory traces. Short-term memory consolidation is defined as a maintenance mechanism that strengthens novel information (Jolicœur & Dell'Acqua, 1998). As demonstrated by previous studies, longer consolidation opportunities (i.e. longer duration devoted to storage) is associated with less decay in memory (e.g. Bayliss, Bogdanovs, & Jarrold, 2015; De Schrijver & Barrouillet, 2017). Thus, Ricker and Vergauwe (2020) pointed out that their long presentation times of storage items [6] might have favored consolidation of the memoranda. Although not discussed by the authors, it is worth noting that other studies using visuo-spatial memoranda with even longer presentation times than what was used by Ricker and Vergauwe (2020) have still produced significant cognitive load effects (e.g. Langerock et al., 2014).

Lastly, Ricker and Vergauwe (2020) suggested that the absence of a cognitive load effect in their study might have resulted from a combination of all the above

[6] The storage episodes in the studies of Ricker and Vergauwe (2020) consisted in displaying items for 400 ms each, followed by a post-perceptual mask and a blank screen each lasting 200 ms.

explanations. In any event, even if the underlying causes of these findings are not fully understood, the results have theoretical ramifications for the TBRS model. As summarized by Barrouillet and Camos (2020), these ramifications probably require to consider different forms of attention or to revisit the predictions of the TBRS depending on the type of recall task considered.

Also, while acknowledging the value of the TBRS model integrating a vast range of working memory functions, we believe it can be further refined in some aspects. The rationale is that in spite of the most recent developments of the model seen in the present section, we believe that the core functioning of the TBRS still merits more attention, in particular regarding the switching process which we believe lacks clarity. One way of addressing this shortcoming is to take advantage of other studies that have integrated a computational architecture to the TBRS framework (e.g.Gauvrit & Mathy, 2018; Lemaire & Portrat, 2018; Oberauer & Lewandowsky, 2011).

1.3 Modeling

Following Baddeley's remarkable progress at conceptualizing working memory, the domain progressively sought to develop models integrating a computational architecture to better predict immediate recall (e.g. Brown et al., 2007; Farrell & Lewandowsky, 2002; Henson, 1999; Lewandowsky & Murdock, 1989; Page & Norris, 1998). Most models then have considered that forgetting in memory is caused by interference of related information, with the exception of the TBRS model that argues in favor of a time-related decay of memory traces in short-term memory.

As we have seen in previous sections, strong empirical evidence supports the TBRS model (e.g. Barrouillet & Camos, 2012, 2009; Barrouillet et al., 2007; Barrouillet, Gavens, Vergauwe, Gaillard, & Camos, 2009; Barrouillet et al., 2011; Plancher & Barrouillet, 2013; Vergauwe, Barrouillet, & Camos, 2010, 2009) and testing cognitive load is now a benchmark (Benchmark 5.2.4. in Oberauer et al., 2018a). However, it has been argued that the model still needs to be implemented to test possible variants (Gauvrit & Mathy, 2018; Oberauer & Lewandowsky, 2011; Portrat & Lemaire, 2015). This observation refers to the limits of the TBRS model, and more generally, to the limits of verbal models. This is not to say that verbal models are not valuable, but rather that computational models allow a theory to go one step further. Since all the parameters of the model have to be considered and defined, the model can be tested with much greater precision. While more precision is clearly a good thing for the advancement of research, the other side of the coin is that the constrained formalism of computation models can sometimes lead to misrepresent the original ideas of the verbal model (Sun, 2008).

The implementation of a verbal model can easily give rise to a whole family of models that predicts different effects (Farrell & Lewandowsky, 2010; Lewandowsky & Farrell, 2010). For instance, the implementation of the phonological loop model of Baddeley (1986) involves no less than 144 variants according to Lewandowsky and Farrell (2010). Regarding the TBRS model, many options can indeed be considered (see Oberauer & Lewandowsky, 2011 for similar arguments). For example, the following clarification should be made: when does the time-related decay start? Does the process of time-related decay begin after the presentation of each item or after the end of the presentation of all items? At what rate does temporal decay occur? Is this rate constant or proportional to the amount of information in memory? Each of these questions leads to the consideration of several alternative options that may have important repercussions on the expected memory performance. The verbal model does not help solving any of these issues, and therefore does not allow to consider clearly how the many potential variants of the model translate quantitatively.

Central to the present thesis is whether switching (between processing and storage) itself consumes time. If yes, switching should be integrated into the equation formalizing cognitive load in the TBRS. If yes, there are still many options to state whether a particular event counts as a relevant switch. Some options are developed in the present work but comparing them was not central in the project. The central objective was to detect the presence of switching costs.

1.3.1 A connectionist model of the TBRS

The TBRS* developed by Oberauer and Lewandowsky (2011) is a two-layer connectionist network model designed to reproduce the activation of the memory traces in complex span tasks. One layer is devoted to the representation of each storage item and the other one handles the representation of serial positions. Through Hebbian learning, a memorandum is encoded by activating the item layer and the position layer associated to it. As time goes by, the strength of the connections follows an exponential growth until reaching an asymptote. This reflects learning. In contrast, an exponential decay is applied to all the items on the list to mimic time-based decay when attention is not focused on storage items. Retrieval is implemented by selecting, recalling, and suppressing the storage item with the highest activation. When items are being refreshed, the same implementation occurs as when items are being retrieved, except that the refreshed storage item is rapidly re-encoded instead of being suppressed. As within the verbal model, refreshing occurs only whenever attention can be focused on the storage items, and it takes place in a cumulative fashion, starting with the first item in the sequence and moving on serially until the end of the sequence. The TBRS* takes into account only one mechanism of maintenance stipulated by the verbal model, viz. refreshing. However, later on Lewandowsky and Oberauer (2015) implemented rehearsal in a simplified version of the TBRS*.

As discussed in the previous section, this implementation of the TBRS model implied to consider many parameters. For instance, Oberauer and Lewandowsky (2011) favored a localist representation of the storage items over a distributed representation of items. This means that each stored item is represented by only one unit in contrast of a pattern of activation across units. Positions, on the other hand, were represented by several units, ensuring thus that neighbouring storage items are more likely to produce serial position errors during recall. Item and order representations are just a few examples of the many modeling choices that had to be considered (see Table 1. of Oberauer & Lewandowsky, 2011 for an exhaustive list of the modeling decisions taken in the TBRS*). Some of these decisions were driven by the theoretical assumptions of the TBRS, but others were motivated by pragmatic reasons. Although it is clear that the TBRS* is a powerful model, it was designed based on no less than 11 modeling choices made by Oberauer and Lewandowsky (2011), which are not specified by the initial model.

1.3.2 A mathematical transcription of the TBRS

A mathematical transcription (TBRS2) of the TBRS model has been developed by Gauvrit and Mathy (2018), which enables the testing of the original model quantitatively. Unlike other computational models (e.g. TBRS* of Lemaire & Portrat, 2018; Oberauer & Lewandowsky, 2011), the TBRS2 only focuses on assessing the cognitive load by implementing strictly the minimal assumptions of the original model. It does not intend to support precise simulations of behavioral data. Its main goal is rather to precisely compute the cognitive load of a task to obtain predictions.

1.3.3 Concept underlying the TBRS2

The TBRS2 conceives that a span equal to k can be simulated whenever the increase of activations is (k -1) times faster than the decrease of activations. Intuitively, if the slopes of activation are considered linear, it does not matter when exactly the activations decrease or increase (as long as their proportion is the same). For instance, let's assume that someone is running up and down the stairs, and that she runs down twice as fast as going up. The jogger can flip directions anytime during her workout but does not stop moving until a stop criterion. The analogy is here only valid to exemplify the simple span task. In the complex span task, one would have to imagine that the person involuntarily slides down whenever the stop criterion is met to mimic decay.

The only predictor of the number of steps climbed is the proportion of time the person goes down before halting. On the long run, the person will go up only if there is at least twice as much time devoted to going up rather than down. This amounts to saying that the person will need to go up at least 2/3 of the time to keep going up. Back to the memory trace interpretation, imagine that activation increases twice as fast as it decreases. In this case, the memory trace of an item will fade away on the long run if it decreases at least twice of the time it increases. The logic is similar to the example above where the protagonist is jogging in stairs faster when going down.

Like in the jogger example, to keep an item in memory, one would need to increase its activation at least $1/3$ of the time (or leaving it decrease no more than $2/3$ of the time). In such a situation, it is possible to maintain up to 3 items concurrently, as long as the activation of each item increases at least $1/3$ of the time. However, with a load of 4 items, it would be impossible to maintain all activations high enough to recover all of the items at the end of a long sequence. To do so, a span equal to 4 would be necessary, which requires that the increase of activation is 3 times faster than the decrease of activation.

To be more realistic, instead of allowing a greater amount of time to reactivate a memory trace, the TBRS considers that the speed of reactivation is faster than the speed of deactivation. In a simple span task, a person for instance will be able to retain 3 items if the speed of reactivation is twice as rapid. It is because the activation is twice as rapid that refreshing can split one third of the time on each of the three items.

Finally, the above explanation shows that a linear function works fine to implement the cognitive load hypothesis posited by the TBRS model, according to which the cognitive load alone determines the span – rather than the specific timeline of free time. However, the fluctuations of activation of the memory traces are supposed to be exponential to be psychologically plausible. The TBRS2 hence bridges the intuitive linear model to the exponential functions by the use of the log of the activation values, since the dynamics of log is linear whenever the odds in the background is exponential. Eventually, the model produces and displays the probabilities of recall which are straightforward to interpret. We develop later how probabilities, logs and odds relate to each other in order to satisfy the characteristics of the TBRS.

1.3.4 Overview of the TBRS2

Figure 1.9 summarizes the general idea of the TBRS2 model. The complex span task is modelled by what Gauvrit and Mathy (2018) call a task function. This function indicates the alternation at a given time between the processing of distractors and the maintenance of to-be-recalled items. In line with the original assumption of the TBRS model, the focus function specifies whether attention is devoted to the processing activities or if it is dedicated to the maintenance of memoranda. In Figure 1.9 the task function consist of two letters to be remembered (i.e. A in red and K in blue) and three distractors represented by black squares. This maintenance occurs via refreshing, which is possible during the free time that is available between the presentation of events (the distractors or the to-be-recalled items). According to the original TBRS model, working memory functions serially, and the refreshing schedule is thought to happen in a cumulative fashion. In other words, during free time, the first item of the list is refreshed first, then the second and so on, until the last item (Barrouillet & Camos, 2012). Accordingly, the TBRS2 refreshes the items in a cumulative manner, but other alternative refreshing schedules are also possible to implement. The refreshing schedule and the task function determine the events of the focus function. In Figure 1.9 we see the alternation of red, blue, and black colors in the focus function. These periods correspond to the storage and maintenance activities of the letters A and K, and the processing of the distractors, respectively. Finally, the focus function shapes the activation of each storage item. Again, in Figure 1.9, the red and blue lines will respectively correspond to the activation levels of the letters A and K. This activation is in the TBRS2 the odds of correct recall of a given item at a given time, and it can be computed from the focus function as soon as decay and refreshment rates are set. As seen, activation increases when the focus function is dedicated to the maintenance of the stored items and decreases otherwise. Importantly in the TBRS2, as in the original model, the only factor influencing the activation of a memory trace is the –inverse– amount of time devoted to refreshment.

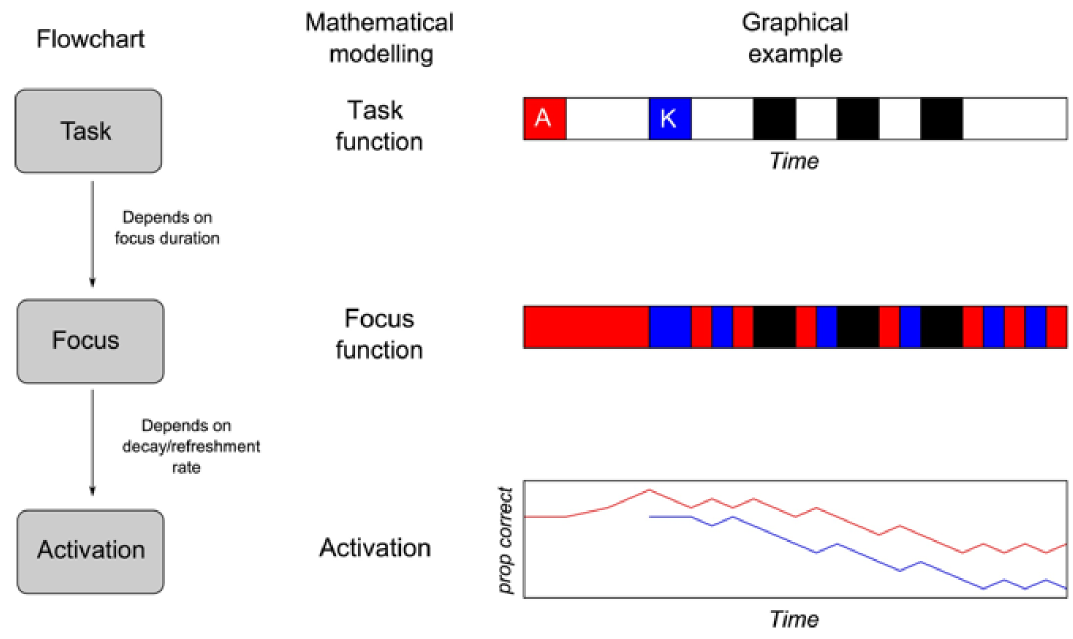


Figure 1.9: Overview of the TBRS2 model. Reprinted from "Mathematical transcription of the 'time-based resource sharing' theory of working memory", by N. Gauvrit and F. Mathy, 2018, Br J Math Stat Psychol, 71: 146-166, p. 149 (https://doi.org/10.1111/bmsp.12112). Copyright © 2017 The British Psychological Society.

Other variants of the TBRS2 exist. The rationale of the model is identical in these variants, but the formula to set the duration and the schedule of refreshing can differ. Regarding the duration, a storage item can either be refreshed for a fixed duration (e.g. 80 ms) or until it reaches a threshold of activation. These variants are called steady and threshold respectively. Regarding the refreshing schedule, we described above the cumulative schedule. However, the TBRS2 can also implement two alternative refreshing orders called next and lowest. According to the next schedule, the model tracks the last refreshed item and after an interruption it starts with the next item in the sequence. In the lowest schedule, refreshing starts with the storage item that has the lowest activation. This seems an optimal strategy to avoid forgetting, but it imposes an extra computational step. Overall, the TBRS2 can implement six different variants of the original model. However, for the sake of clarity and following previous work (e.g. Oberauer & Lewandowsky, 2011), in this dissertation we will essentially conduct simulations using a cumulative schedule and a fixed duration of refreshing.

1.3.5 Probabilities, odds & log-odds

The TBRS2 implements the original verbal model using probabilities, odds, and log-odds. All computations are made in log-odds, which allows the TBRS2 to establish the activations of the to-be-recalled items. As we discussed earlier, these activations are expressed in odds of success, however, all the outputs and the plots of the simulations that we will present show these activations in terms of probabilities of correct recall for better readability. The probabilities are a sigmoid function of log-odds. Although, probabilities, odds, and log-odds provide essentially similar information, the respective use of either one is in the TBRS2 determined by mathematical constraints. To avoid any confusion stemming from the terminology, let us recall the fundamentals before describing further the model.

In order to better grasp the relation between the notions and the respective outcomes they provide, the following helps visualize a range of values in one scale and convert them to the other scales[7].

For instance, if the probability of remembering correctly an item is 0.2, then the corresponding odds of the outcome are:

$$odds = \frac{0.2}{0.8} = 0.25$$

Likewise, the corresponding log-odds are:

$$ln\left(\frac{0.2}{0.8}\right) = -1.3863$$

The initial probability can be retrieved from the odds thanks to:

$$\frac{odds}{1 + odds} = \frac{0.25}{1.25} = 0.2$$

[7]The example was taken from the tutorial https://www.montana.edu/rotella/documents/502/Prob$_{o}dds_{l}og-odds.pdf$.

And finally, probability can be computed from log-odds as:

$$\frac{exp(ln(odds))}{1 + exp(ln(odds))} = \frac{exp(-1.3683)}{1 + exp(-1.3683)} = \frac{0.25}{1.25} = 0.2$$

Recall that within the TBRS2, activations are computed in log-odds, so essentially for a given sequence of memoranda we obtain the probability of correct recall thanks to:

$$\frac{exp(ln(activation))}{1 + exp(ln(activation))}$$

As seen in Table1.1, log-odds ranging from -5 to 5, and odds ranging from 0.01 to 148.41 correspond to probabilities that range from above 0 to close to 1. Figure 1.10 shows the relation between odds and log-odds, probability and odds and probability and log-odds. These values are only examples, as odds range within $[0; \infty[$.

Table 1.1: Conversion of Log-odds, Odds and probabilities.

$Log-odds$	$Odds$	$Probability$
-5.00	0.01	0.01
-4.00	0.02	0.02
-3.00	0.05	0.05
-2.00	0.14	0.12
-1.00	0.37	0.27
0.00	1.00	0.50
1.00	2.72	0.73
2.00	7.39	0.88
3.00	20.09	0.95
4.00	54.60	0.98
5.00	148.41	0.99

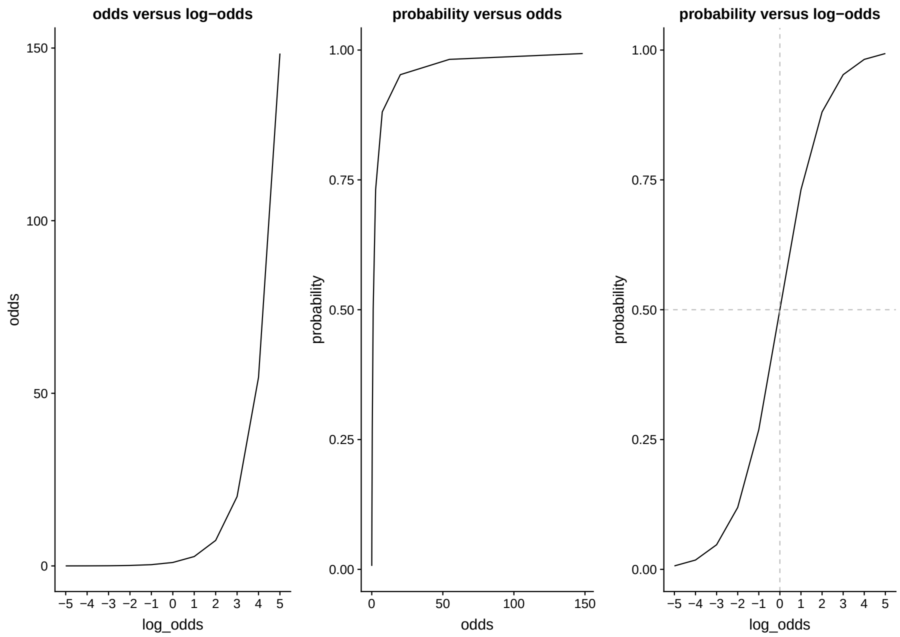


Figure 1.10: Conversions of odds *vs.* log-odds, probability *vs.* odds and probability *vs.* log-odds.

TBRS2 in summary

The model can be freely simulated at:
`https://mathematicalpsychology.shinyapps.io/tbrs/`.

- The *decay rate* d corresponds to how fast the memory traces decay. Based on previous empirical publications Mueller and Krawitz (2009), a recommended value is $d = 1.4$.
- The *refreshing rate* corresponds to how fast a memory trace is reinforced. Based on the average simple span 7 (Miller, 1956), $r = 9$ can be deduced from $\frac{r+1}{d} = 7$.
- The *duration* is the time devoted to refreshing a memory trace. It can be set to 80 ms following Oberauer and Lewandowsky (2011).
- The *baseline* can be derived from a Brown-Peterson task to simulate a loss of memory content of 50% after 3.5 seconds of distraction.
- The switch sw is an optional parameter representing the time needed to switch from the concurrent task to the memory task.

1.3.6 TBRS2 explained

The TBRS2 model comes with 5 parameters:

Decay rate. In all decay-based models, the decay rate is an essential parameter, as it estimates the loss of the memory trace. In the TBRS2, the decay rate (d) corresponds to the decrease in log-odds of correct recall (memory traces) when no refreshing occurs during a second. In the field of psychology, decay has often been described in terms of duration. Previous research has estimated that a memory trace vanishes between 150 ms and 20 s, depending on the type of information. Most of these estimates range between 5 and 10 s (e.g. Byrnes & Wingfield, 1979; Kieras, Meyer, Mueller, & Seymour, 1998; Mueller & Krawitz, 2009; Sams, Hari, Rif, & Knuutila, 1993).

Based on these previous empirical publications, in particular Mueller and Krawitz (2009), we suggest a value of $d = 1.4$. According to the authors, the probability of correct recall decreases from .8 to .2 in about 2 seconds; that is, to fit the function of the TBRS2, the log-odds go from $\log(4)$ to $-\log(4)$ in 2 seconds, thus losing $log(4) \approx 1.4$ points per second. This number may not seem very intuitive as it is not expressed in seconds, but in log-odds lost per second. In other words, d is roughly a multiplicative factor that says by how much the activation is divided in a second (more exactly: it is divided by exp(d)).

Figure 1.11 illustrates the decay rate of the TBRS2 of only one item in memory when no refreshing opportunity is allowed. The top of the plot displays the structure of the task, with the events depicted horizontally as a function of time. The black line represents the successive events of the concurrent task. The line entitled Focus represents the letter that is at the center of attention. We can see that when attention is fully captured by the secondary task, memory traces cannot be refreshed. Thus, the decrease in activation of the storage item is exponential. After 4.5 seconds (right side of the frame), the level of activation predicts the probability of recalling each of the items.

Refreshing rate. Another crucial parameter in the TBRS2 model is the refreshing rate, which is the theoretical gain in the activation of a memory trace during a second devoted to refreshing. Akin to the decay rate, the refreshing rate is expressed in terms of log-odds of correct recall. Under the assumptions of the original model, the decay rate and the refreshment rate are bound to be closely linked to each other. Indeed, in the TBRS2, because the ratio $\frac{r+1}{d}$ corresponds to the simple span, it is supposed to revolve around 7 to follow Miller (1956). With this in mind, the rate needs to be set to $r = 9$.

Duration. The duration is the time devoted to refreshing one item, generally set to 80 ms, following previous empirical findings. Previous research has indeed estimated that the average refreshment rate is between 40 ms to and 80 ms (e.g. Jarrold, Tam, Baddeley, & Harvey, 2011; Oberauer & Lewandowsky, 2011; Vergauwe et al., 2014; Vergauwe & Cowan, 2015).

Figure 1.12 represents how refreshing occurs in the TBRS2 when several items are to be recalled. Once again, the top of the plot displays the structure of the task, with the events depicted horizontally as a function of time. Each color represents a to-be-recalled item. In this example, there is no black line(s) on the task function, meaning that no concurrent task was displayed. From the Focus line, we can see the different to-be-recalled items entering the center of attention. At every free delay opportunity, we can see that previously presented items are being refreshed successively in a cumulative fashion. After 15 seconds, the level of activation predicts the probability of recalling each item. One can note that the first and last items represented respectively by the red and fuchsia colors are expected to be best remembered, which recalls the famous serial position function.

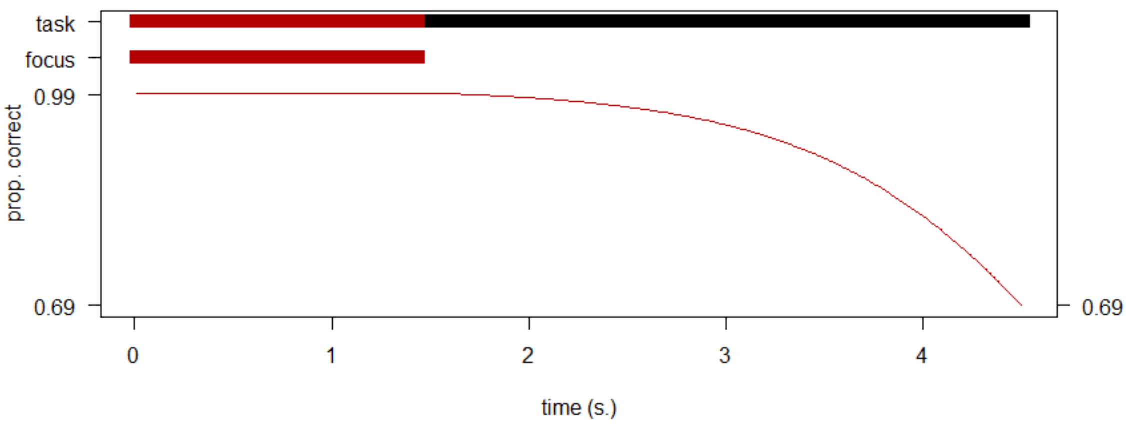


Figure 1.11: Illustration of the decay rate of the TBRS2 for one to-be-recalled item presented for 1500 ms followed by a continuous series of processing events lasting 3000 ms. The TBRS2 predicts a probability of correct recall of 0.6899745 after 4.5s. This simulation involved the following parameters: start type = "first", type = steady, baseline = 5, duration = 8, $d = 1.4$, $r = 9$, $sw = 0$.

Baseline. The baseline is the log-odds of correct recall of an item that is currently presented. In the interest of simplification, the TBRS2 considers that whenever a storage item is presented for the first time, its activation will depend on a constant baseline value as long as it is displayed. Although this is not explicitly described in the original model, this does not undermine in any way the generality of the TBRS2 as long as the presentation duration of the storage events are constant across the memoranda. Mueller and Krawitz (2009) showed that (in their log-normal model) the probability of recall equals $.5$ after a delay of 4 seconds. With a value of $.5$ to recall an item after a distraction of 4 seconds, the odd of recalling the item is $(.5/.5 = 1)$; thus, the log-odds$(1) = 0$. Since the TBRS is based on a linear decay of the log-odds, this amounts to saying that the baseline needs to be approximately equal to $4 \times 1.4 = 5.6$. To retrieve the probability of recall at the baseline, we compute $exp(5.6)/(1+exp(5.6) = .99$. If we simulate the model by taking a baseline of 5.6, with $d = 1.4$ and $r = 9$, the probability of recall returned by the model is effectively $.5$. Because in the famous Brown-Peterson paradigm the loss of the memory content has rather been estimated to be around 50% after 3 seconds, leading to a baseline of 4.2, we use an intermediate rounded value equal to 5 in most our simulations. As seen from Figure 1.13, the TBRS2 predicts a probability of correct recall just under 1 after one second of attentional capture with a baseline of 5. This probability decreases progressively to $.90$ and $.69$ after two and three seconds, and hits 0.35 after four seconds of attentional capture.

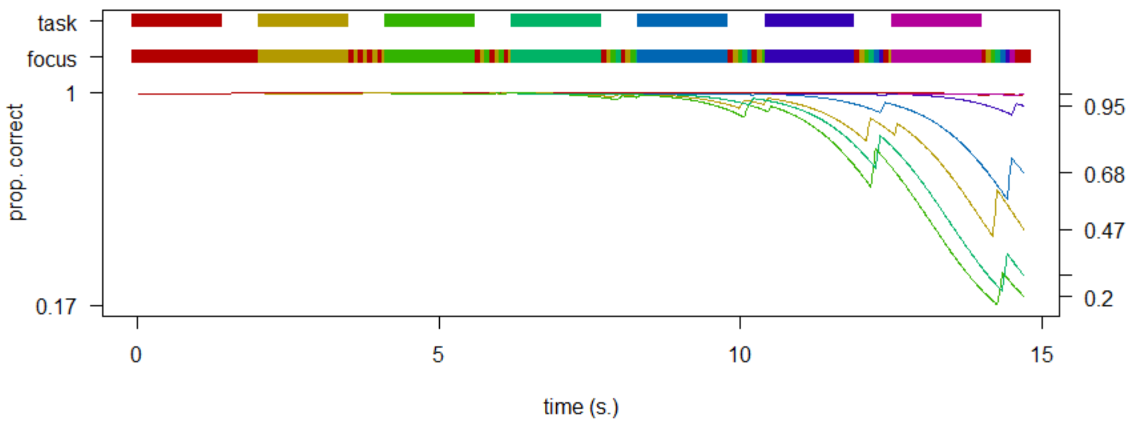


Figure 1.12: Illustration of the refreshing cumulative schedule and the rates of activation for seven to-be-recalled items presented for 1500 ms each, followed each by a free delay of 600 ms during which refreshing can occur. The TBRS2 predicts a probability of correct recall of 0.9920295, 0.4650571, 0.2042403, 0.2857736, 0.6848174, 0.9470498, 0.9932538 for the seven consecutive items. This simulation involved the following parameters: start type = "first", type = steady, baseline = 5, duration = 8, $d = 1.4$, $r = 9$, $sw = 0$

Switch. The switch sw is an optional parameter linked to executive functions corresponding to the period of time required to switch from the concurrent task to the refreshing of the memoranda. This parameter was developed in the interest of testing further the potential influence of switching costs in complex span tasks.

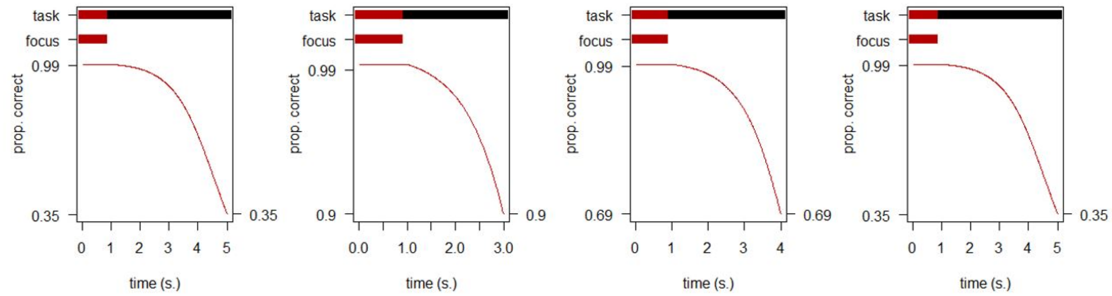
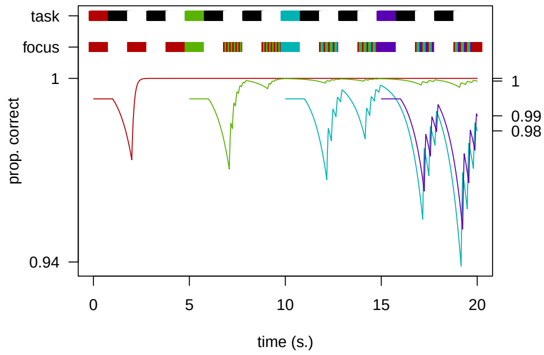


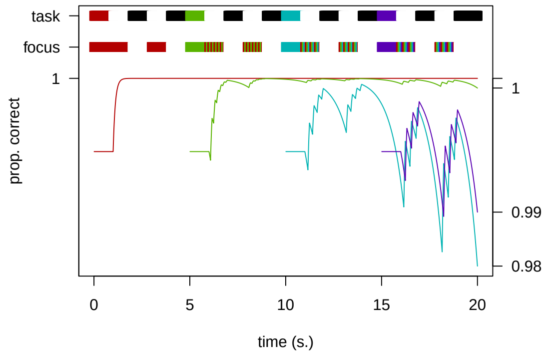
Figure 1.13: Illustration of the simulations of the TBRS2 with a baseline fixed at 5. The plots (from the right to the left) shows the probability of correct recall for one letter after one, two, three, and four seconds of attentional capture. The TBRS2 predicts a probability of correct recall of 0.97, 0.90, 0.69, 0.35 for the letter after one, two, three, and four seconds respectively. This simulation involved the following parameters: start type = "first", type = steady, baseline = 5, duration $= 8$, $d = 1.4$, $r = 9$, $sw = 0$

Overall, the TBRS2 is a transparent implementation of the TBRS model. Figure 1.14 illustrates this point, and shows four simulations involving four storage items interspersed with distractors and free delays dispatched in various orders (with the same cognitive load). In the caption of each subfigure, the task function is translated in what we will call in the rest of this dissertation a processing pattern or schedule. In each pattern, L stands for a letter to recall, 0 for a delay of free time, and 1 for a distractor. The task function and its respective processing pattern provides the same information in a different format, the latter one is simply more useful in the body of the text. Importantly, the simulations in Figure 1.14 show that the TBRS2 predicts the same probability of recall at constant cognitive load unconditionally of the processing schedule used. As seen, each plot shows slightly different probabilities of correct recall as a function of time, however, the end result is the same in each case, as long as each memorandum is refreshed equally.

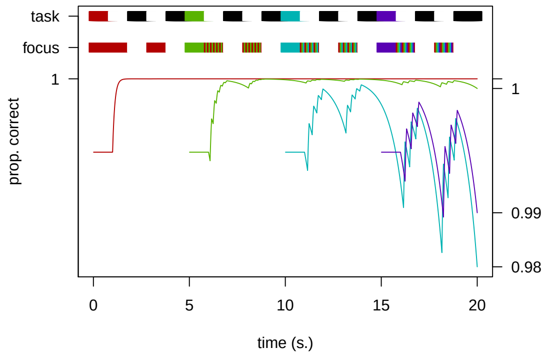
Since the main purpose of the TBRS2 is to explore the activation levels of storage items over time, it does not rely on many parameters or on a specific architecture. For this particular reason, the TBRS2 is a useful mathematical tool to investigate the potential effect of switching on memory performance, without making any additional suppositions that is not already discussed in the TBRS verbal model. The next section will briefly discuss the switching literature.



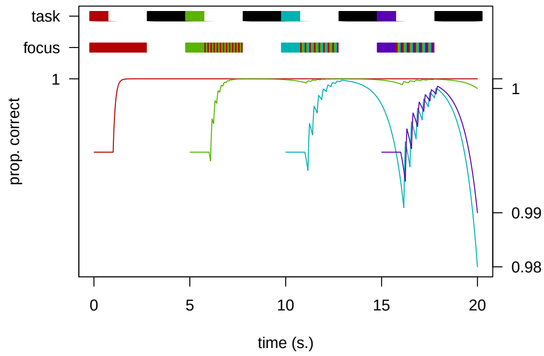
(a) Processing pattern: L10101010



(b) Processing pattern: L01010101



(c) Processing pattern: L11001100



(d) Processing pattern: L00110011

Figure 1.14: Simulations of the TBRS2 for four items to be remembered, interspersed with different processing patterns that correspond to a same cognitive load. The TBRS2 predicts for the four different tasks the exact same probability of correct recall for the sequence in mind: 0.9699707. *Note:* L stands for storage items, 0 for a delay of free time, and 1 for a distractor. This simulation involved the following parameters: start type = "first", type = steady, baseline $= 5$, duration $= 8$, $d = 1.4$, $r = 9$, $sw = 0$. The duration of all events (i.e. storage, distractor, and free delay) was set at 1 s.

1.4 Switching

The executive role of working memory is generally viewed as allowing individuals to regulate the contents of their memory flexibly based upon information that needs to be kept active or inhibited (Ellis, 2002; Miyake & Shah, 1999). This is in line with the TBRS model, which conceives working memory as an executive system. It is therefore no coincidence that the complex span task has been used extensively to test this model.

As mentioned above, the task devised by Daneman and Carpenter (1980) was originally thought to study individual differences in reading comprehension by alternating times of processing and storage. Various versions of the task have proven useful to estimate working memory capacity, and these estimates have been shown to be closely correlated to intelligence (Colom, Flores-Mendoza, & Rebollo, 2003; Kane et al., 2004; Unsworth & Engle, 2006; Unsworth, Redick, Heitz, Broadway, & Engle, 2009). To manipulate both to-be-remembered information and distractors, participants in complex span tasks are asked to recall a set of items in correct order while they are required to engage in processing activities unrelated to the memory items (Aben, Stapert, & Blokland, 2012; Unsworth & Engle, 2007b). This task has been most often associated to the latent statistical construct of working memory, while the simple span task (which only requires to recall a set of items in correct order) is most often associated with the short-term memory construct.

In complex span tasks, because storage requests and processing events alternate regularly, a core idea of the TBRS is that this alternation requires task switching. By definition, switching between tasks refers to the ability to shift from several tasks. This process is sequential and it involves disengaging attention from an irrelevant past task while actively engaging in the subsequent task (Miyake et al., 2000). While this definition is in line with what is thought to occur during a complex span task according to the TBRS model, the mechanism of switching has been little explored by the authors. The idea developed in the present work is that switching might be more time-consuming than originally thought by the creators of the TBRS model, and as such it might affect how cognitive load is conceived.

1.4.1 Switching paradigms

Although the complex span task fits the definition of task switching since participants are required to alternate between storage and processing activities sequentially (according to the TBRS model), switching has typically been studied through paradigms of its own. Most often the experimental tasks make use of a same set of stimuli for both tasks (Kiesel et al., 2010), and the nature of the tasks do not necessarily involve any memory load. For instance, the participant would be required to switch between adding 3 to a series of digits and subtracting 3 to another set of digits. Sometimes, the first task is easier than the second one, this allows to study whether it is more difficult to return to the most complex task or to inhibit it. Whatever the task-specific characteristics used in an experimental design, there can be a cost for the participant to deal with two tasks. The theories in this domain are thus easily applicable to the complex span task that requires switching from the memory task to the concurrent task.

Since the seminal work of Jersild (1927), numerous studies have shown that task switching is typically associated with higher response latencies and higher error rates compared to conditions where a single task is repeated (see Monsell, 2003), and this cost is found even when the switches are predictable (Rogers & Monsell, 1995). Task switching was first explored with the use of lists (e.g. Allport, Styles, & Hsieh, 1994; Jersild, 1927), where performance or time needed to execute pure lists are compared to mixed lists (i.e. task AAA or BBB *vs.* task ABABAB). For example, the pure list could correspond to a task where participants are asked to add 3 on every digit presented (task A) or to subtract 5 from them (task B). The mixed list would then be composed of an alternation between additions and subtractions. In list procedures, the switch cost corresponds to the difference between the time needed to execute a pure list minus the time needed to perform the mixed list (Vandierendonck, Liefooghe, & Verbruggen, 2010). More complicated subcomponents of switching can be identified (e.g. mixed-list costs *vs.* alternation costs) to better identify the preparatory component and the residual component of switching, but these more elaborated aspects will not be fully developed in the present thesis.

The list procedure has some drawbacks, as other determinants may explain in part the observed switching cost. As argued by Rogers and Monsell (1995), working memory may be one of these factors, as participants need to keep track of the alternation between the two tasks. According to Rubin and Meiran (2005), another factor that may contribute to switching costs in mixed lists is related to an increased amount of interference between task sets (i.e. the cognitive processes involved in the tasks[8]) in situations where the same kind of stimuli is used in both tasks. Although these findings highlight the limitations of list procedures, they also allow to shed light on the processes involved in what is referred to as the mixed-list cost, which is typically found in this paradigm. The mixed-list cost relates to the observation that reaction times in task-repetition trials are shorter in single-task blocks compared to a mixed-task block (Vandierendonck et al., 2010), for instance when the transition AA is performed in AAAAAAA versus in AABBAABB.

To avoid these limitations, other procedures are nowadays preferred (Kiesel et al., 2010). Among these, the most commonly used methods are the alternative runs and the cuing procedure (e.g. Altmann, 2007; Rogers & Monsell, 1995; Meiran, 1996). In alternative runs, participants complete only mixed lists, and the first task is performed a certain number of times before switching to the second task (i.e. AABBAABB). The switch is thus predictable, since participants are informed that every n trial, they are required to switch tasks. In order to decrease the working memory load related to the monitoring of the task switch, the next task is determined by spatial cues of the current trial. For instance, targets presented on the right side of the screen are associated to task A, whereas the left side is associated to task B as shown in the panel A of Figure 1.15. Another variant of this procedure consists of using unpredictable trials as seen from panel B of Figure 1.15, in which the position of the arrow at 12 o'clock will systematically indicate a task switch (Vandierendonck et al., 2010). Because there is no longer the need to introduce a pure list, task-switching costs are measured within a single block by comparing performance of switch trials to repetition trials (Kiesel et al., 2010).

Although the predictable alternative runs method eliminates some of the disadvantages of the list procedures, some authors argue that a possible confound may overestimate the switch cost (Vandierendonck et al., 2010). This confound relates to the phenomenon of restart cost (also known as a first-trial cost), in which

[8]The notion of task set is often ill-defined and include a wide range of task parameters. Rogers and Monsell (1995, p.208) illustrate the concept of task set through the following example in which participants are requested to choose optional responses as fast as possible: "To perform one of these simple reactive tasks, the participant must somehow chain together and configure an appropriate set of processes linking sensory analysis to motor output. These processes must include categorization of sensory input with respect to a particular attribute or set of attributes; mapping the attribute's value by means of a decision criterion to one of a predetermined set of response categories; and execution of the motor response used to signal that response category." This example gives an overview of how complex the apparent simple notion of switching can be.

participants show worse performance for the first trial compared to the second one in task-repetition runs. The restart cost highlights the processes implemented specifically on the first trial of a run, and it is not determined by whether a switch occurs or not (Altmann, 2007). Many studies have shown that the restart cost occurs when the response-stimulus interval (RSI) is different for the first and the second trial (e.g. Altmann & Gray, 2002; Gopher, Armony, & Greenshpan, 2000; Waszak, Hommel, & Allport, 2003). The role of item timing has been put forward as potential cause to explain this phenomenon. According to Altmann (2007), although the alternative runs paradigm uses a constant RSI between trials, there is a confound between the position of the run and switching costs. Switch-trials always occur in the first position of the run, whereas the task-repetition trials necessarily occur in the second position of the run. Hence, switching costs could be overestimated in predictable alternative runs designs, because they include the restart cost associated to the first position of the trial. Another flaw of this paradigm is the fact that this method does not allow to dissociate task preparation from the time devoted to the dissipation of the earlier task set. Both processes supposedly happen during the RSI. Failure to dissociate these aspects from one another makes this paradigm uninformative to understand what happens during the RSI, and to determine to what proportion task preparation and task-set dissipation contribute to the switch cost (Vandierendonck et al., 2010).

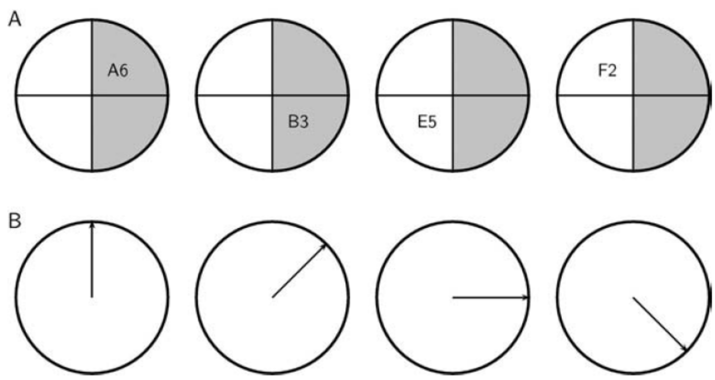


Figure 1.15: Illustration of a clockwise presentation of the stimuli in the alternating-runs procedure. Figure A demonstrates a predictable sequence, in which the task switch is to be performed when the vertical midline is crossed. The gray and white areas represent two distinctive tasks. Figure B illustrates an unpredictable sequence. The arrow designates the task to perform. A task switch is to be performed when the arrow indicates 12 o'clock. Reprinted from "Task switching: interplay of reconfiguration and interference control", by A. Vandierendonck, B. Liefooghe and F. Verbruggen, 2010, Psychonomic Bulletin Review, 136(4), 602 (https://doi.org/10.1037/a0019791). Copyright © 2010, American Psychological Association.

The cuing procedure enables to address some of the shortcomings of the alternative runs paradigm, as the time course is better controlled for in this method. This procedure involves a task cue that informs the participant which task is to be performed. Figure 1.16 illustrates all three paradigms discussed above. As in the alternative runs method, the cuing procedure calculates the switch cost by comparing task-switch trials with task-repetition trials in an unpredictable sequence of events. Task preparation may be examined by manipulating the cue-stimulus interval (CSI), while the time needed to dissipate the previous task set can be examined by varying the RSI (Vandierendonck et al., 2010). Controlling these parameters is important, since CSI is often pointed out as a variable influencing the magnitude of the switch cost (e.g. Monsell, 2003; Logan, 2003). In addition to the switch cost found in switch-trials, the cuing procedure allows to demonstrate that usually response times decrease further when task-repetition trials occur several times in a row (Kiesel et al., 2010).

A potential concern about the cuing procedure is the amount of information provided by the cue. Aside the fact that cues can be arbitrary or transparent (i.e. cue X *vs.* cue $parity$ for a parity judgment task), and that greater switch costs are usually associated to arbitrary cues, another matter to take into consideration is that switching is always associated with a change of cue whereas task repetition is not (Vandierendonck et al., 2010). To avoid this problem, Mayr and Kliegl (2003) suggest the information-reduction paradigm, in which several cues are used for the same task. The first situation is the typical task-repetition trial where the task and the cue remain the same across two different trials. The second situation is the typical switch condition, in which a different cue will signal a task switch. Finally, the third situation, called by the authors a cue-switch condition, repeats the same task but with a different cue (i.e. a non-informative cue). The results of Mayr and Kliegl (2003) from a series of experiments using this modified cue paradigm showed an important cue-switch cost, meaning that performance can be hindered by a change of cue, even if the task remains the same.

Overall, there are many versions of switching procedures in the literature, each with its advantages and limits (Kiesel et al., 2010; Vandierendonck et al., 2010). Switching costs can be investigated by comparing performance across different blocks or trials. Paradigms such as the list procedure compare single-task blocks (i.e. pure lists) to task-switching blocks (i.e. mixed lists), which highlights global costs (also known as general costs). In contrast, paradigms such as the alternative runs or the cuing procedure allow to demonstrate local costs (also called specific costs) by comparing repetition trials to task-switching trials (e.g. Ellefson, Shapiro, & Chater, 2006; Kiesel et al., 2010). From one study to another there is important variations in terms of the magnitude of switching costs that are observed. Conclusions across procedures are ill-advised, as aside the divergence of methods used to investigate switching costs there is also many variations in the tasks used (e.g. arithmetic, categorization, spatial judgment, word naming, reaction choice tasks) (Vandierendonck et al., 2010).

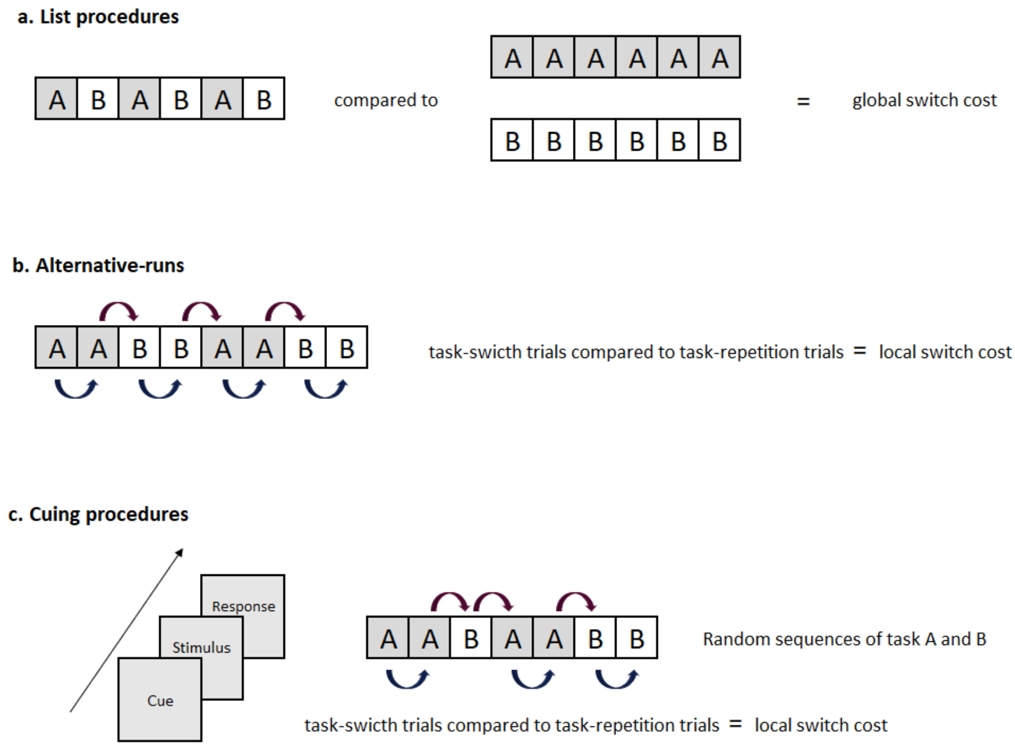


Figure 1.16: Illustration of some switching procedures. Figure a. represents the list procedures. Figure b. represents the alternative-runs procedure. Figure c. illustrates the cuing procedure. Reproduced from "Control and interference in task switchin—A review", by A. Kiesel, M. Steinhauser, M. Wendt, M. Falkenstein, K. Jost, A. Philipp, I. Koch, 2015, Psychonomic Bulletin Review, 17(1), 851 (https://doi.org/10.1037/a0019842). Copyright © 2010, American Psychological Association.

Task switching in summary

- Task switching was first explored using list procedures allowing to explore **global switch costs** which correspond to the difference of performance between a single-task block and a double-task block.

- Other procedures such as the alternative runs and the cuing procedure are preferred nowadays. These methods allow to demonstrate a **local switch cost** which is calculated by comparing task-repetition trials to task-switch trials.

- The **mixed-list cost** refers to shorter reaction times in task-repetition trials in single-task blocks compared to a double-task blocks.

- The **restart cost** (or the first-trial cost) relates to the finding that performance in the first trial is worse compared to subsequent trials in task-repetition runs.

- The **cue-switch cost** relates to the finding that performance is hindered by a cue switch (i.e. a lure), even if no task switch is required subsequently.

1.4.2 Explaining the switch cost

Although task switching is an active field of study, the processes responsible for switch costs are still debated (e.g. Arrington & Logan, 2004; Vandierendonck et al., 2010). Generally, two alternative hypotheses are put forward to explain switching costs, viz. the reconfiguration or the interference account.

According to the reconfiguration account, switching costs result from active processes that enables to load a task set. The concept of task set is central in this theory; it covers the cognitive processes involved in the tasks, including the mental representation of the task goals and the corresponding stimulus-response mapping (Kiesel et al., 2010). To paraphrase a helpful analogy from Monsell (2003), reconfiguration may be thought as a "mental gear change" that is time consuming. This is an action in itself. For instance, these active processes can be identified as goal shifting and rule activation, which implies updating working memory and retrieving task-set information from long-term memory (Rubinstein, Meyer, & Evans, 2001). Other processes such as attention shifting between different stimuli, which allows their identification and the retrieval of appropriate responses from long-term memory and their execution have also been put forward to account for these active processes (Logan & Gordon, 2001).

According to an alternative point of view, the interference account, switch costs are caused by more passive processes that are set in motion by the product of interference between several tasks to be executed (Allport et al., 1994; Wylie & Allport, 2000). Simply put, the change of task in itself would prompt executive processes to resolve interference between the previous and the present task. This interference would lead to switch costs, whether reconfiguration of the task set occurs or not. This interference is referred by Allport et al. (1994) as task-set inertia, which relates to the idea that activation from the previous task set lingers on and impede performance of the present task. Other sources of interference have been discovered, but the main idea of this account is that switch costs emanate from the resolution of interference (Vandierendonck et al., 2010).

In line with the interference account is the finding of asymmetrical switch costs by Allport et al. (1994), in which the tasks to perform are unequal in terms of difficulty[9], such as naming a color instead of reading the colored word in the Stroop task (Stroop, 1935). The execution of the color-naming task is usually considered to be less automatic than the word-naming task. The asymmetrical switching cost refers to the finding that the cost associated to switching is more important when the shift occurs from a difficult task to an easy one, compared to the other way around. This was interpreted by the authors as evidence in favor of the task-inertia hypothesis, as the switch cost seems determined by the task participants are switching from and not as a function of the task they are about to perform. Initially, proactive interference was discussed as the mechanism involved in this effect, in the sense that the previous task set could affect the execution of the upcoming task. However, further inquiries specified that this proactive interference could also emerge “as a form of long-term negative priming” from previously learned stimulus-response bindings between a set of stimulus characteristics and responses (Wylie & Allport, 2000). Aside the exact mechanism of interference in play in this phenomenon, what matters here is that the amount of inhibition needed to block

[9] In the literature, different terms have been used to distinguish the tasks that produce an asymmetrical switch cost. For instance, Allport et al. (1994) and Wylie and Allport (2000) use the terms dominant *vs.* non dominant. In contrast Yeung and Monsell (2003b), characterize the tasks as strong *vs.* weak. Ellefson et al. (2006) refer to the task as difficult *vs.* simple. Whatever the label chosen, the weaker task is harder to execute because it is less familiar, less practiced, more complex, etc.

the previous task set is determined by the difficulty of the previous task, and this concurs with the time needed to overcome this inhibition. Hence, an easier task will demand less task-set inhibition, which takes less time to resolve.

Although, the asymmetric switch cost has been replicated many times with different types of tasks and populations (e.g. Ellefson et al., 2006; Filippi, Karaminis, & Thomas, 2014; Meuter & Allport, 1999; Yeung & Monsell, 2003a), it is not systematically found (e.g. Costa & Santesteban, 2004; Hübner, Kluwe, Luna-Rodriguez, & Peters, 2004; Monsell, Yeung, & Azuma, 2000). Moreover, asymmetric switch costs may be reversed, and according to Yeung and Monsell (2003b) it should occur only when the two tasks highly interfere with each other. This is the case especially in tasks that share some stimulus features and response sets, in which the selection of the relevant stimulus characteristic and appropriate response are thus harder.

Another finding that call into question the generalization of the asymmetric switch cost and its interpretation in terms of the interference hypothesis alone is the observation of asymmetrical costs without the need to actually switch to another task (Bryck & Mayr, 2008; Gopher et al., 2000). This result referred as a restart cost by Bryck and Mayr (2008) was produced in their studies with the alternative runs method, in which a pause of variable length (500 or 5000 ms) was introduced after each pair of trials (i.e. AA–pause-AA–pause-BB–pause-BB). Their results showed that an asymmetrical cost was present when a long delay was inserted between trials not only in task-switch trials (i.e. A-B), but also in task-repetition trials (i.e. A-A or B-B), hence indicating a possible loss of the task set in working memory. These results are discussed by the authors in terms of a long-term memory retrieval hypothesis. A premise of this view is that participants need to recover the relevant task-set information in long-term memory after task switching or when this information has simply been lost over time, as it was the case in the experiments introducing a long delay between trials. In order to explain asymmetrical switch costs typically found in tasks that are unequal in terms of difficulty, the authors postulate that encoding the task set associated to a difficult task requires additional attentional control in comparison to a simpler task. This additional attention control leads to a stronger activation of the task set of the difficult task in long-term memory, which in turn makes the difficult task more likely to provoke interference during retrieval attempts.

Overall, this rapid review of the switching literature provides a glance at the complexity of the field. Without doubt, the domain outlines a rich conceptual background to account for complex span task performance. Since switching will be investigated in the present work only through the complex span task paradigm, several other findings pertaining to very specific paradigms in the switching literature have not been discussed here (but for a review see Kiesel et al., 2010; Vandierendonck et al., 2010). Finally, before describing the core of this dissertation, we will in the next section present a brief summary of the studies bridging together the working memory and switching literature.

1.4.3 Working memory & switching

Working memory and switching are known to be closely related (Miyake et al., 2000). For instance, some studies have focused on the role of verbal working memory in task switching. In this regard, the work of Goschke (2000) is interesting, as it showed that verbalization can benefit task-switch performance. Effectively, participants who were required to verbalize the task before their execution showed reduced switching costs, and this was not observed when the same tasks were executed while saying an irrelevant word (e.g. "Monday" or "Tuesday"). More direct evidence in favor of the role of verbal working memory in task switching was brought by Baddeley et al. (2001). In their study, the authors used a list procedure in which participants where required –under articulatory suppression– to complete additions or subtractions in pure lists and to alternate between these two types of operations in mixed lists. The results showed as expected a mixed-list cost. But more interestingly, this cost was increased under concurrent articulatory suppression, whether the concurrent articulatory task involved executive processes or not (i.e. explicit category switching *vs.* irrelevant word repetition, respectively "January, Monday, February Tuesday"... *vs.* "the, the, the..."). In a similar vein, Emerson and Miyake (2003) showed in a series of nine experiments also using the list procedure, that articulatory suppression increased the magnitude of switch costs substantially (from 59 to 150% depending on the experiment). Overall, these findings along with many other studies corroborate the idea that verbalization contributes to task switching by enabling the maintenance of the action program of the upcoming task (Bryck & Mayr, 2005; Liefooghe, Vandierendonck, Muyllaert, Verbruggen, & Vanneste, 2005; Miyake, Emerson, Padilla, & Ahn, 2004; Saeki & Saito, 2004, 2009). However, as discussed by Emerson and Miyake (2003), the role of verbal working memory in task switching seems limited to the situations demanding endogenous control of the task switch. In other words, when external cues are provided for the task switch, the negative impact of articulatory suppression on switch costs is greatly reduced. Note, however, that these studies based on articulatory suppression are a proof by contradiction (reductio ad impossibile), which is not totally satisfying since articulatory suppression can also distract attention. A less indirect proof of the role of verbal working memory in switching was provided by Laurent et al. (2016) using surface laryngeal electromyography, which contrary to previous studies enabled participants to verbalize the tasks freely. The authors showed that inner speech was recruited more often for tasks demanding endogenous control.

Beyond the exploration of the role verbalization in task switching, the relation between working memory and switching has received little attention. The task-span procedure developed by Logan (2004, 2006) was, however, an attempt to bridge together research on working memory and switching. The task-span is composed of two steps in which the participants are first asked to memorize a list of task-names of various length. The task-names refer to three different tasks: a magnitude task (e.g. is the digit smaller or bigger than 5?), a parity judgment task (e.g. is the digit presented odd or even?) and a form judgment task (e.g. is the digit presented numerically or is it written?). During the second step of the task, digits were displayed one-by-one in their Arabic or written form, and the participants were required to execute the previously memorized corresponding tasks to the stimuli. Hence, the task-span procedure required working memory to both maintain the to-be-executed tasks and keep track of the ongoing progress while completing the task itself. Additionally, the procedure involved switching since several tasks were to be completed within a list of tasks. The results showed no difference between the task-span and a typical memory span in which the participants recalled only task-names (i.e. with no requirement to switch between several tasks). Consequently, Logan (2004) argued that there is no trade-off between switching and storage capacity, indicating that task switching implicates processes outside of working memory. In line with Logan's findings, other work also found no influence of memory load on task-switching performance (Kiesel, Wendt, & Peters, 2007; Liefooghe, Barrouillet, Vandierendonck, & Camos, 2008).

The opposite relation, the effect of task switching on recall performance, was investigated by Liefooghe et al. (2008). The authors conducted several experiments in which they inserted task switching in a complex span task paradigm. In other words, instead of having one concurrent task in a complex span task, the participants were required to switch between two digit judgment tasks (magnitude and parity judgment task) while memorizing letters. The results of the first experiments showed that switching reduces recall performance as a function of the number of task switch to perform, showing clear evidence in favor of detrimental effects of switching costs in working memory. According to the TBRS model, a reduced recall performance should be proportional to the time during which attention is not available to refresh the memory traces. In order to test the prediction of the TBRS model, Liefooghe et al. (2008) conducted a fourth experiment in which they examined working memory performance in three different experimental conditions: (1) low number of task switch to execute between the parity and magnitude judgment tasks with visually degraded stimuli; (2) low number of task switch to execute with no degradation of the stimuli; (3) high number of task switch to execute with no degradation of the stimuli. The idea was that the use of visually degraded stimuli would occupy attention to the same extent that the high task-switch condition while not influencing the magnitude of the switch cost. The results confirmed the predictions of the TBRS model and showed that when the amount of attentional capture is equated, the number of task switch to execute no longer produces significant differences on memory performance.

To summarize, previous studies investigating the interaction between working memory and task switching have found contradictory results. However, when time parameters are strictly controlled for, the work of Liefooghe et al. (2008) brings evidence in favor of switch costs in working memory. According to this account, switching has a detrimental effect on working memory because it captures attention away from the maintenance activities. In this regard, switching has the same detrimental effect on memory performance as any other processing activity does. Although this finding clarifies the relation between task switching and working memory, the work of Liefooghe et al. (2008) does not address whether switching between tasks is fundamentally akin to switching between storage and processing activities. Therefore, it remains unclear whether switching between storage and processing (and the other way around) has also a detrimental effect on working memory. The present dissertation is an attempt to fill in this gap in the literature.

However, within the TBRS model it has not been clearly posited that the storage of items can be conceived as task that is associated to a specific task set. We consider this to be the case, on the grounds that storage items require to be maintained either through rehearsal or refreshing, in the same way that processing a distractor is mapped to a task set because it requires its own mental operations (e.g. preparing to judge whether a digit is even or odd). This is also in line with the position of Vandierendonck, Szmalec, Deschuyteneer, and Depoorter (2007, p. 252), who argued about the operation-span task that "(...) it requires the participants to achieve two task goals. On the other hand, it is necessary to remember all the words in the presentation order. On the other hand, it is necessary to perform well on the calculation task. Hence, the participants must configure two task sets that are kept active during the entire test procedure."

Positing that a complex span task involves two task sets, the following questions need to be answered: (1) how can we manipulate experimentally switching between storage and processing activities in complex span tasks? (2) how do we define a switch between storage and processing activities in a complex span paradigm? (3) Among the many variants of dual task paradigms that exist in the literature, which type of task would be best suited to explore switch costs between storage and processing activities within a complex span task?

How can we experimentally manipulate switches between the storage and processing activities in complex span tasks?

Knowing that complex span tasks are usually composed of a regular alternation between storage and processing subtasks, how can we experimentally manipulate the number of switches between these activities in such a task? We introduced a first glance at a potential experimental setup which could allow the number of switches to be manipulated in the section describing the TBRS2 model. Recall that in Figure 1.14 we varied what we called the processing patterns within a complex span task. Following this same logic, we suggest that by manipulating the order of the distractors and the free delays between the to-be-recalled items, the number of switches to execute between storage and processing activities will also vary in consequence. For instance, we hypothesize that within the following sequence 10101010 and 11001100 the amount of switches between storage and processing activities vary. This leads to our second question.

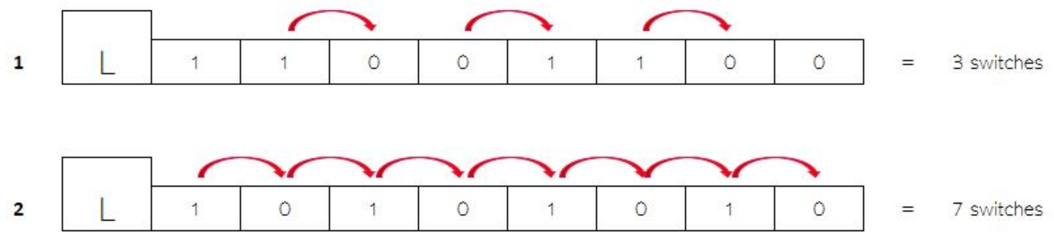
How do we define a switch between the storage and processing activities in a complex span paradigm?

The study of Liefooghe et al. (2008) described earlier allowed the authors of the TBRS model to make the conjecture that switching occupies the central bottleneck, which in turn postpones refreshing and leads to the degradation of the memory traces. Therefore, we suggest that the idea of switch costs between storage and processing activities taxing complex span task performance befits the TBRS model. However, there are several possibilities for conceptualizing the switches between the two main functions of working memory. In the following section, we will consider two different accounts: the TBRS2 and our interpretation of the original TBRS model.

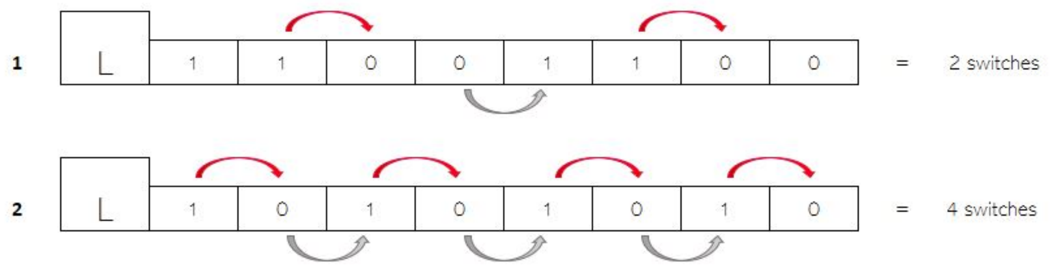
a) The TBRS2 account

Figure 1.17 represents two ways to implement switches within the TBRS2 model (i.e. a parameter could vary the cost, starting with the value zero). The most intuitive way to consider switches between storage and processing activities is to visualize a switch whenever a different type of event is presented during the task. Essentially this means that the number of switches is dictated by the structure of the task. Figure 1.17a represents this possibility. As seen, every time the task displays a series of distractors followed by a free delay or the other way around, a switch is expected to take place. Hence, in Example #1 we can see that when the processing and the free delays are grouped together, this amounts to saying that there should be three switches in the sequence. In contrast, Example #2 in Figure 1.17a represents a more typical complex span task structure made of a regular alternation between events, in which there should be seven switches.

On the other hand, within the TBRS(2) model working memory performance is conditioned by the cognitive load of a given task. The cognitive load of a task encompasses the proportion of time during which attention is not directed towards the maintenance of the memory traces. Therefore, it could be argued that the detrimental effect of switching from the storage activity to the processing activity is in fact already captured within the cognitive load equation of the model. In contrast, the switch costs occurring from processing to storage are not captured by the cognitive load. To assert that switch costs between the two main functions of working memory are indeed partially included in the cognitive load function, we assume that switching between storage and processing (and the other way around) is akin to any other attention-demanding mechanism (such as switching between tasks). We also assume that the cost associated to a switch should not be conditional on the type of task about to occur. In other words, switching from storage or from processing should incur the same cognitive cost, or at least a same threshold of detectable switch cost. Figure 1.17b shows the number of switches



(a) Switches between storage and processing (and the other way around).



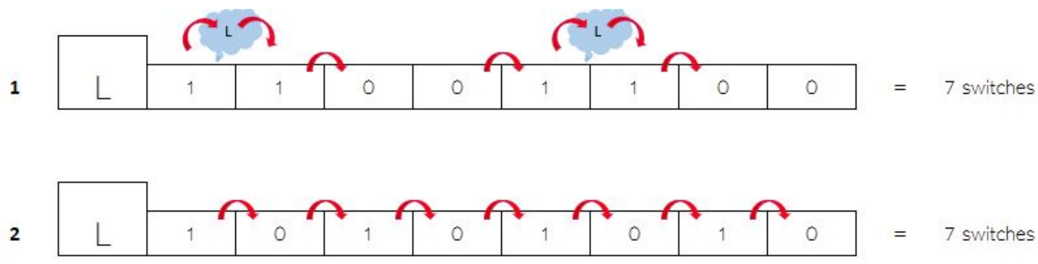
(b) Switches between storage and processing (and the other way around). The cognitive load function accounts for the attentional capture of the switch from storage to processing.

Figure 1.17: Figure (a) represents our TBRS2 conceptualization of a switch between storage and processing in a complex span task. The red arrows specify the switches for each processing schedule. Figure (b) represents an alternative view considering that the cognitive load function in the model already captures the cost associated to switching from storage to processing The red arrows specify the switch costs not included within the cognitive load function. The grey arrows represents the costs that could be considered as already captured by cognitive load. L stands for a storage item, 1 for a distractor and 0 for a free delay.

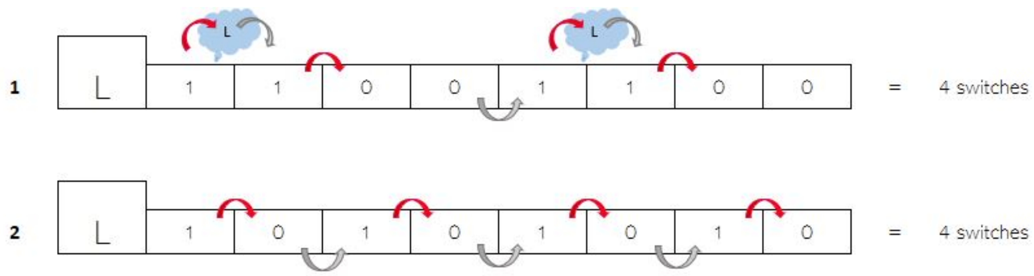
according to this view. The grey and red arrows specify the switch costs that are respectively accounted for or not by the cognitive load function. Overall, we can see that although the number of switches differs from the previous option (Figure 1.17a), the overall prediction of the TBRS2 model remain unchanged, as there are still fewer unaccounted switches in the processing schedule of Example #1 than in the processing schedule of Example #2. Hence, according to the TBRS2 views we expect impaired complex span task performance as a function of the number of switches occurring in a trial.

b) The original TBRS account

We believe that Figure 1.18 represents conceptions closer to the original model. So far, the previous accounts have determined the number of switches in a complex span task by taking into consideration solely the structure of the concurrent task. However, within the TBRS model switching between storage and processing is thought to happen *continuously*. This continuous switching process is also paired with a high-speed refreshing mechanism (about 50 ms per refreshed item) which allows to reactivate the degraded memory traces. Subsequently, switching could occur before and after *each* processing step, but also during the concurrent activity by taking short pauses in order to refresh the memoranda (Barrouillet & Camos, 2014b; Camos, 2017). In this case, the structure of the task becomes irrelevant, if refreshing can happen between two successive distractors. As seen from Figure 1.18a, the number of switches in the two processing schedules of Example #1 and #2 is now equated when refreshing between two processing steps is also taken into account. It is worth noting that the TBRS model conceives more frequent refreshing episodes than what is actually represented in the Figure 1.18a. For instance, refreshing could also occur during a processing step. For the sake of clarity, we depicted only the refreshing episodes occurring between two consecutive distractors to emphasize the difference between the original TBRS and the TBRS2 models. The conceptualization of the switches of the original TBRS model would essentially mean that the attentional capture associated to the execution of the processing task and the potential switch cost are not distinguishable from one another, since every processing step is coupled with a switch. From this point of view, it could be stated that the switch costs between storage and processing in our processing schedules can only be tacitly assumed at best, but it is not possible to quantify them. In other words, according to this conception of the switches the core hypothesis of switch costs between storage and processing (and the other way around) predicts a null effect in the experiment that we will develop.



(a) Every processing step is coupled with a switch.



(b) Every processing step is coupled with a switch. The cognitive load function accounts for the attentional capture of the switch from storage to processing.

Figure 1.18: Figure (a) represents our understanding of the original TBRS model views on switches between the two main functions of working memory. Figure (b) represents the same view while adding several assumptions not explicitly stated by the original model. The red arrows specify the switch costs not accounted for in the cognitive load function of the original model. The grey arrow represents the costs that could be considered as captured by the cognitive load in the original model. The blue bubble stands for a refreshing episode of the to-be-recalled letters. For the sake of clarity only a single refreshing episode occurring between two consecutive processing steps is represented. Other refreshing episodes are not depicted. L stands for a storage item, 1 for a distractor and 0 for a free delay.

Figure 1.18b represents the views of the original TBRS model, while adding assumptions not explicitly stated by the model. Essentially, this account considers that the switches from storage to processing are captured by the cognitive load function. This leaves only the switches occurring from the opposite direction unaccounted for (i.e. from processing to storage). The same assumptions described above need to be made as in Figure 1.17b. As seen from the Figure 1.18b, taking into consideration the cognitive load factor does not change the prediction that the two processing schedules remain equal in terms of the number of switches they include in Example #1 and #2. In sum, the variations of the structure of the task that we plan to test should not have any influence on working memory performance according to this view.

The endorsed view of this thesis

This work will endorse the first TBRS2 option (represented in Figure 1.17a), which considers switches as a result of the task structure. We believe that this view provides a neutral position to start examining the effect of switching on complex span task performance. This option does not add any additional assumptions to our hypothesis. Additionally, it enables us to provide a testable hypothesis to study in a general way the influence of subtle variations of the structure of the complex span task. This last point is of particular interest for examining untested predictions of the original model because according to the TBRS2 model such variations could influence recall independently of the cognitive load. The present thesis thus posits that the structure of a complex span task should matter, in particular if switch costs corresponding to our TBRS2 view occur.

Which type of task would be best suited to explore switch costs between storage and processing activities within a complex span task?

We have established in our introduction that there are many variants of the complex span task. These variants concern the type of concurrent task (e.g. operation-span task, symmetry span, reading-digit span task, parity judgment task). In other words, the nature of the concurrent task and the processes that it requires can greatly vary from one complex task to another. We also have established that all these tasks have in common the fact that they alternate storage and processing activities in working memory. In that respect, we argue that all these tasks should be valid candidates to investigate the effect of switching costs within complex span tasks. However, we believe that there is a straightforward opposition among the tasks used in the literature that remains unexplored. The opposition is about whether the task begins with a storage or a processing episode. When the task starts with a storage episode, it simply implies that a memory load is added to the task from the beginning, whereas this is not the case when the task begins with the concurrent task. It remains an open question whether this subtle variation of the structure of the task has any implications in regard to working memory capacity, and by extension to our main question of interest in the present work. We propose to examine this question in more depth in our first study before pursuing further our inquiry on the potential switch costs in working memory complex span tasks.

1.5 Chapter description

As we have established, working memory refers to a complex cognitive system that allows the temporary storage and processing of a limited amount of information. According to the TBRS model, working memory functions by rapidly switching between processing and storage activities. The complex span tasks generally used to evaluate working memory require this rapid alternation by using a primary memory task interspersed with a secondary processing task designed to direct attention away. The goal of the present dissertation was to manipulate the structure of the complex span task to explore as of yet untested predictions from the TBRS. A key idea is that the TBRS model, thus far, considers switching as a by-product of serial processing but switching has not yet been studied thoroughly itself within the canonical complex span task. The present dissertation therefore posits that switching costs should be taken into account to refine estimates of the cognitive load associated with a concurrent task.

In Chapter 2, we will discuss the fact that although the complex span task is without doubt the most commonly used measure of working memory capacity, the structure of this task varies from one study to another. Surprisingly, it has not been questioned yet whether these variants could influence working memory capacity or their predictive power of fluid-intelligence. Hence, our first experiment is an attempt to remedy this neglected aspect of the literature. As we will see, processing events within a complex span task may influence the *estimates* of working memory capacity. We will discuss these results in the light of the TBRS framework.

In Chapter 3, we examine further whether the structure of the processing task can alter working memory performance more than expected in the original TBRS model. More specifically, this Chapter focuses on the influence of switching costs, which we argue have gone unnoticed in past and current experimental studies. Switching costs are induced by any kind of multitasking situations, and our objective was to propose that they could be better implemented into the TBRS model. For this purpose, we ran a series of experiments using complex span tasks in which the number of switches was manipulated with different rhythmic patterns within the concurrent task. With this goal in mind, we expected to give a different account of the cognitive load, which is central in the TBRS model. As we will discover, the results showed that these patterns can slightly influence working memory performance in some cases but not in others.

In Chapter 4 and 5, we address the limits of our manipulation of switches as the rhythmic patterns used in our secondary tasks may have introduced a confound variable. Some of the patterns containing fewer alternations between storage and processing introduced longer delays of free time, possibly enhancing the consolidation of the memoranda. Therefore, we decided to explore this confound with another series of experiments using rhythmic patterns designed to test specifically the switching cost and the consolidation hypotheses.

Finally, the last Chapter of this dissertation will provide a general discussion about our main question of interest regarding switch costs in complex span task performance. Here, we will review the central premises of our work and attempt to draw a comprehensive account of our findings. We conclude that the TBRS model in its current implementation remains the most parsimonious way to account for performance in complex span tasks, regardless of subtle variations of performance that can be introduced when manipulating the structure of the task.

Complex span tasks & fluid intelligence

2

The complex span tasks used to assess working memory capacity are predictive of many aspects of higher-order cognition However, the structures of these tasks vary from one study to another, and it has never been called into question whether these variations could influence their predictive power. Previous research has exclusively used two types of complex span task structures, either those based on alternating processing-storage patterns (e.g. the operation span task of Unsworth, Heitz, Schrock, & Engle, 2005) or alternating storage-processing patterns (e.g. the operation span task used by Barrouillet & Camos, 2001). The difference between these two types of tasks relies on whether the task begins with a processing episode that captures attention (e.g. solving mathematical operations, reading digits aloud) or by a storage episode involving the maintenance of presented items (e.g. letters, words, digits, shapes). We believe that the potential effect of the various structures of complex span tasks should not be overlooked for two main reasons.

First, complex span tasks are extensively used in correlational studies, for instance, in examining the relationship between the span and fluid intelligence. Depending on the material used in complex span tasks, variations from moderate to strong correlations have been found (e.g. Kane et al., 2004; Kanerva & Kalakoski, 2016; Lucidi, Loaiza, Camos, & Barrouillet, 2014; Unsworth et al., 2009). Our hypothesis is that the structure of the task could also influence the strength of the correlations. It is unclear why complex span tasks are so predictive of higher-order cognition (Conway et al., 2003; Kane et al., 2005; Unsworth et al., 2009). According to the dual-component model of Unsworth and Engle (2007b, 2007a), attention control and secondary memory are crucial mechanisms for both working memory and fluid abilities. Attention control enables the keeping of relevant information accessible despite interference, and secondary memory allows access to long-term memory, and they both contribute distinctively to higher-order abilities (Unsworth & Spillers, 2010). In a recent article, Engle (2018) specifies that attention could operate differently on working memory and higher order abilities. In working memory tasks, attention control primarily enables the maintenance of information while other irrelevant thoughts are being suppressed. In a reasoning task (e.g. Raven's

Progressive Matrices Raven, 1962; Matrix Reasoning of the WAIS-IV Wechsler, 2011) the main role of attention control is to allow participants to disengage from information once it turns out to be irrelevant. However, maintenance processes are also involved in reasoning tasks, but in a lesser degree. Although not completely incompatible with this previous account, the Time-Based Resource-Sharing model (TBRS, see Barrouillet et al., 2004) supposes that the relation between working memory capacity and intelligence is best explained by the fact that both constructs require switching between processing and storage within the time constraints of the task at hand. In that sense, general purpose resource sharing is essential, rather than the combined effect of attentional control and secondary memory (Barrouillet et al., 2007; Barrouillet & Camos, 2007; Barrouillet, Lepine, & Camos, 2008; Lucidi et al., 2014). Studying the structure of the span task could therefore shed light on the mechanisms underlying the correlations with intelligence. Second, but relating to the first point, we think that manipulating the structure of the task could help decide between models when they are implemented in various versions (e.g.Gauvrit & Mathy, 2018; Lemaire & Portrat, 2018; Oberauer & Lewandowsky, 2011). It is possible that one of the structures weights more on a given process (i.e. storage vs processing) at different positions (i.e. potentially impacting primacy and recency effects; see Unsworth & Engle, 2006), and the structural variations could thus impact empirical correlations and the fit of alternative models. The comparison of the fit of alternative models is beyond the scope of the present study. Instead we evaluate the possible impact of the structure of the task on working memory performance, and its relation to fluid intelligence.

2.1 Subtle variations of complex span tasks

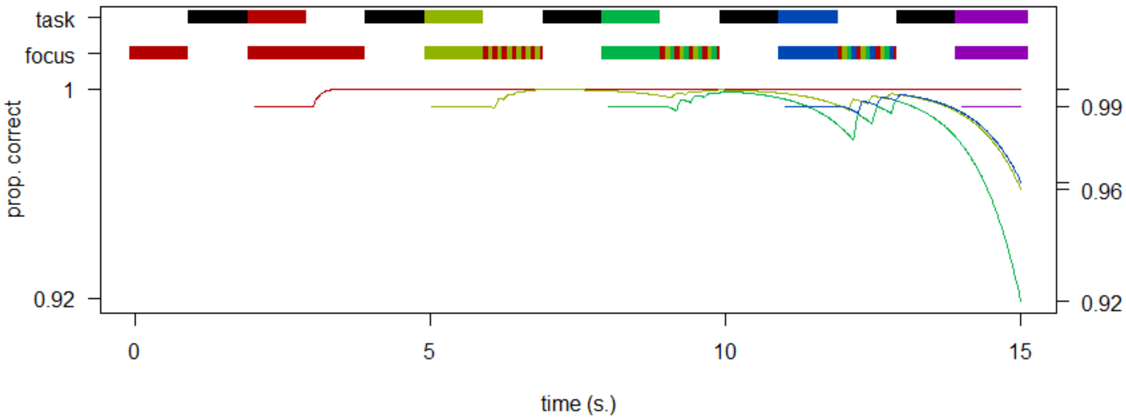
The operation span task of Unsworth et al. (2005) is an example of a complex span task based on a processing-storage repeating pattern, where the participant is invited to solve mathematical operations that are interspersed with items to remember. Crucially in this type of task, the task begins with a processing episode, and the task ends with a storage episode. The operation span task has also been used by Barrouillet and Camos (2001), but the structure of the task is reversed, as it starts with the storage episode and ends with the processing episode. This difference warrants attention because the complex span tasks may be the most popular task for studying working memory (Aben et al., 2012; Barrouillet & Camos, 2012; Barrouillet et al., 2011; Colom, Rebollo, Abad, & Shih, 2006; Dunlosky & Kane, 2007; Gathercole, 1999; Friedman & Miyake, 2005; Lewandowsky, Oberauer, Yang, & Ecker, 2010; Mathy, Chekaf, & Cowan, 2018; Miyake et al., 2000; Oberauer, 2009; Oberauer et al., 2018a, 2012; Ricker, Vergauwe, Hinrichs, Blume, & Cowan, 2015; Schmiedek, Hildebrandt, Lövdén, Wilhelm, & Lindenberger, 2009; Wang, Ren, Li, & Schweizer, 2015), and this task offers interesting theoretical insight when its temporal structure is manipulated (Bailey, 2012; Loaiza & Camos, 2016; McCabe, 2010). However, the choice of one of the two structures is generally adopted without clear justification. For instance, Stone & Towse (2015, p.3) mention that "the *traditional* method of administering complex span tasks such as the operation span task involves using a processing-storage order of phases rather than storage-processing (...). This method is rather curious as the processing task serves the purpose of adding to the cognitive demands of storage by requiring processing of a task while trying to store stimuli."

As explained above, the executive account of working memory (e.g. Engle, 2002, 2018; Engle, Kane, & Tuholski, 1999) posits that working memory capacity is essentially defined by the ability to maintain controlled attention (in order to allow storage processes to take place) despite interference. Thus, the structural variations of the processing items (at least for medium and longer list lengths) should not influence working memory capacity because the demands in controlled attention should be equal in a processing-storage and a storage-processing complex span task. For instance, the processing-storage version of the task should not necessarily benefit the last item because the requirement to recall all of the items in order could sufficiently equalize the maintenance of information requirement.

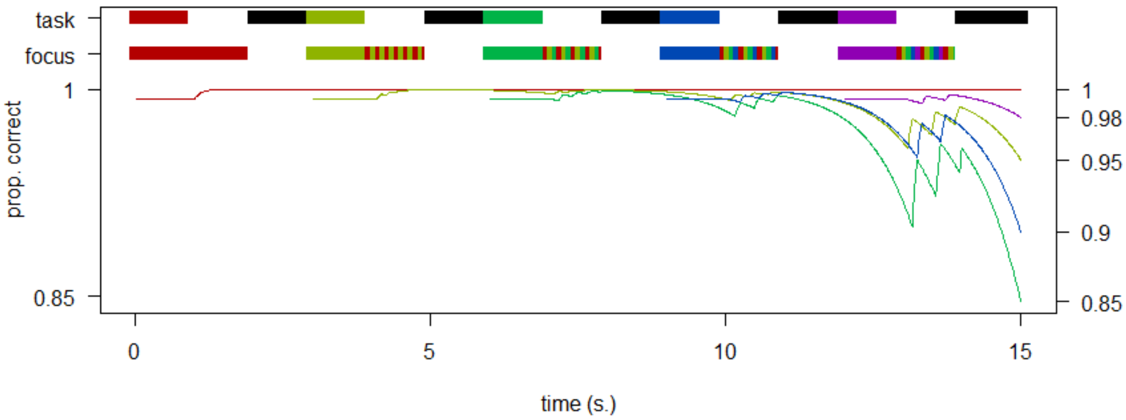
In contrast, according to the TBRS model (Barrouillet et al., 2004) working memory spans are dependent of the cognitive load (CL) of the task at hand. As we already discussed in the first chapter of this thesis, cognitive load is defined as the proportion of time during which a given task occupies attention during memory retention, preventing thus maintenance processes to occur. Several studies have found a linear relationship between the complex span and the cognitive load of the task (e.g. Barrouillet et al., 2004, 2007, 2009, 2011). Thus, the greater the proportion of time during which attention is fully occupied (i.e. the higher CL), the lower the memory performance. Recall also, that experimentally, cognitive load is manipulated by varying either the duration of the attentional capture, the number of processing steps and/or by changing the total amount of time available for the participant to perform the processing task. For instance, increasing the number of operations to solve in a complex span task while keeping the other parameters constant increases the CL of the task at hand. Following this rationale, the processing-storage tasks should overestimate working memory capacity by minimizing the processing demand, as the first item of the task (i.e. solving a mathematical operation) is memory load free. Consequently, this first mathematical operation and its duration should be insignificant to the cognitive load of the task, whereas the same operation in the storage-processing version of the task should augment the cognitive load. In sum, the TBRS account can more easily predict differences between the two types of tasks than the above executive account.

The mathematical transcription of the verbal model (TBRS2) allows to test this prediction with more precision. Figure 2.1 shows the predictions of the TBRS2 when the structure of the complex span task was manipulated. The top of each plot displays the structure of the task, with the events depicted horizontally as a function of time. Each new colored event on the top row of each plot represents a new to-be-remembered item. All of the black events represent the successive events of the concurrent task. The line entitled 'Focus' represents the letters that are successively at the center of attention. We can see that free time allows the different memory traces to be refreshed alternatively. The curves below show the level of activation of each of the to-be-remembered items. After 15 seconds (right side of the frame), the level of activation predicts the probability of recalling each of the items.

As seen from the numbers that are reported in the figure's note, the probability of the correct recall of the letters in the processing-storage type of task Figure 2.1a is on average higher. These differences seem low using the default parameters, but these numbers could be better differentiated by using different sets of parameters. However, crucially, our prediction is that because the processing-storage type is always granted with a lower cognitive load, participants' spans should be higher. This is intuitive, as the last processing episode of the storage-processing type of task is detrimental to the recall process. In that respect, we can also expect that the more demanding storage-processing type of tasks are better predictors of higher-order cognition.



(a) Predictions of the TBRS2 for an alternating processing-storage type of task. The probability of correct recall for the successive five letters are: 0.9999766, 0.961432, 0.9186392, 0.9642913, 0.9933071.



(b) Predictions of the TBRS2 for the storage-processing condition: 0.9999683, 0.9484365, 0.8460564, 0.8965995, 0.9794474.

Figure 2.1: Predictions of the TBRS2 for the processing-storage and the storage-processing conditions. Both simulations involved a sequence of five letters and the parameters were set as follow: Duration 80ms, baseline 5, $d = 1.4$, $r = 9$.

2.2 Experiment 1

The paper reporting the data of this study was published by Debraise, Gauvrit, & Mathy in Psychonomic Bulletin & Review (2021)[1]. In the present experiment, the participants were submitted to both processing-storage and storage-processing conditions of an operation span task. After completing a simple span task (i.e. a span task without any concurrent events) and both the conditions of the complex operation span task (counterbalanced between the participants), the participants then performed a reasoning test (Matrix Reasoning of the WAIS-IV Wechsler, 2011). The two versions of the operation span task (OSpan) were adapted from Unsworth et al. (2005) and Barrouillet and Camos (2001), respectively. However, the timing of both conditions was based on that of Unsworth et al. (2005). Note that we here only focus on performance for the complex span tasks and the reasoning test, as the simple span task was intended to serve as a pretest for another study. We therefore consider that the simple span task was used herein as a warmup.

2.2.1 Method

Participants. Two hundred two young adults (164 females, 38 males; mean age = 21.52 years, $SD = 3.9$) participated in this experiment. Most of the participants were second-year students of Université Côte d'Azur and received partial course credit for participating. All participants were tested in a quiet room at the university. The duration of the experiment was approximately one hour.

Regarding power, we followed recommendation of Schönbrodt and Perugini (2013) to reach a sufficiently large sample of participants to obtain stabilized correlation values. After collecting data on $N = 83$ participants, we obtained the following estimates: $r = .409$ for the processing-storage condition and $r = .367$ for the storage-processing condition. Based on these values corresponding to the rounded estimate $\rho = .4$, we followed the authors' recommendation to collect data based on at least $N = 150$ participants for a corridor of stability of width $w = .15$ and a level of confidence equal to 95%[2]. Also, we made sure we could obtain a significant

[1]Experiment 1 was not formally preregistered; De-identified data for experiment 1 is posted at https://osf.io/rh6qb/. The materials used in these studies are widely available. The authors wish to thank the action editor of Psychonomic Bulletin Review, Gene Brewer for the appropriate advice on power analysis.

[2]The range of sample sizes was between $N = 342$ for $w = .10$ and $N = 84$ for $w = .20$, but the larger value exceeded our pool of participants who could obtain course credits in exchange of

difference between the two estimates, with plausible correlation values. Based on the lowest value $r = .367$, we verified that we had reached a sufficient sample size using the test for two dependent Pearson correlations in G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). The computation showed that the second correlation should be at least .12 higher than the smaller one with $N = 180$ participants to reach a power of .80, which seemed a plausible difference to be observed between different measures of the span. Finally, we also computed the minimal sample size to reach a power of .80 in our mediation analysis using the function runGitHub in R, which indicated a sample size of $N = 170$.

Material and Procedure. The simple span task consisted of three trials for each list of 3 to 9 letters randomly chosen without replacement from the following set of consonants: F, H, J, K, L, P, Q, R, B, S, V, X. Each trial started with a fixation cross displayed for 500 ms centered on the screen, followed by the to-be-remembered letters, each presented for 1000 ms. At the end of each trial, a matrix with all the consonants appeared at the center of the screen. The participants were then invited to recall the letters in the correct order by clicking on the letters. A feedback of the number of letters correctly recalled for the current set was presented on the screen before a new trial started. Five practice trials of a set size of two letters were administered before the simple span task began. The set of to-be-remembered consonants, their duration and the recall phase were identical in the complex span tasks that followed. Before the first complex span task, a training block of mathematical operations took place where the participant had to solve 16 mathematical operations as fast as possible. These trials began with a cross presented for 500 ms followed by a chosen mathematical operation (e.g. $(2 \times 6) - 4 =?$). After solving the operation mentally, the participant was instructed to click on the mouse in order to pass to the next screen where a digit (e.g. 8) and the words 'VRAI' and 'FAUX' ('TRUE' and 'FALSE' in French, respectively) were displayed. The participant was requested to click on the correct answer, again as quickly as possible. In this example, the participant would have to click on the 'VRAI' box in order to be correct. After each operation was solved, the percentage of correctly solved mathematical operations was updated and displayed on the right corner of the screen. The participants were instructed that they had to reach a 85% success rate so that their results could be included in the study. The aim of this mathematical operation training was to familiarize the participants with the concurrent task but also to measure the average individual's response times (RT) to tailor the main task, as explained later. Similar mathematical operations (i.e.

their participation the year we finished conducting the study in 2020.

same kind of difficulty) were used in the complex span tasks. Both complex span tasks had similar structure in terms of cognitive load, distractors, and memoranda. Both complex span tasks used list lengths varying from 3 to 7 consonants and three trials per list length. The 15 trials (5 lengths $\times$ 3 trials) within each complex span task were randomized for each participant. Before each complex span task, the participants practiced with two trials of length two (using either the processing-storage order or the storage-processing order).

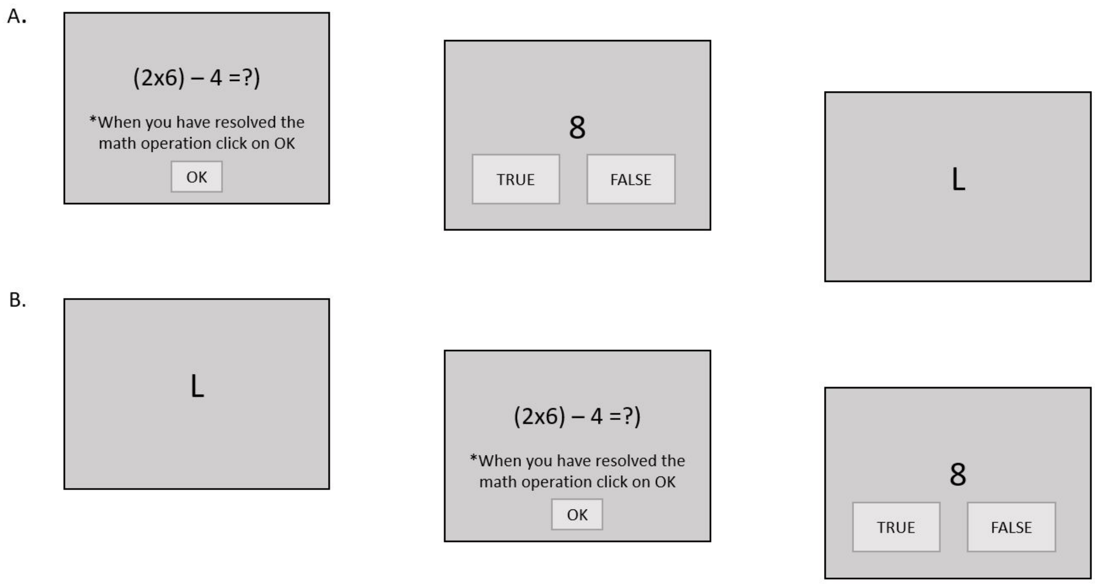


Figure 2.2: Illustration of two types of structure in the operation span task. A = Processing-Storage task; B = Storage-Processing task. *Note*. Instructions were in French in the experiment.

The only variation between the two types of task was the order of the processing and storage events. The processing-storage task (Unsworth et al., 2005) consisted of presenting the mathematical operations before the consonants. Each trial began with a fixation cross presented for 500 ms followed by a mathematical operation. After solving the mathematical operation, a letter was presented on the screen for 1000 ms, followed by a new operation and so on, until the end of the trial and the recall phase. The participants were instructed that their available time to solve the mathematical operation would be limited to pace the task. As in (Unsworth et al., 2005), we used the distribution of the individual RTs obtained during the practice trials to define the available duration based on their average RTs plus 2.5 SD. When the participants did not solve the operation within the available duration, the task resumed as if the participant had actually responded (the program skipped the screen displaying the response options 'VRAI' and 'FAUX'), and the program presented the next event (i.e. a consonant). Then, the trial continued its cycle as it was planned, but the missed mathematical operation was counted as an error. The storage-processing condition (Barrouillet & Camos, 2001) consisted of presenting the mathematical consonants before the operations. The procedure was identical to the processing-storage condition except that each trial began with a consonant. After both complex span tasks were completed, the reasoning matrix subtest of the WAIS-IV was conducted. This test is a nonverbal reasoning task consisting of 26 items including three practice items. For each item, a matrix of colored figures is presented, and the participant must find the missing figure among five response options. This test was not performed for all participants, because some of our participants had previously obtained the sufficient amount of course credits.

2.2.2 Results

The results from two participants were excluded from the analyses as they interrupted the experiment. No other participant was excluded from the analyses, although some did not achieve an 85% average solving rate on the mathematical operation. According to Unsworth et al. (2009), this exclusion criteria is unnecessary, as the processing accuracy is positively correlated to the storage accuracy. Thus, excluding participants with a low processing score may lead to a bias where low-span individuals are also excluded. Overall, the processing accuracy was at 89% ($SD = 0.08$) and 87% ($SD = 0.08$) for the processing-storage and storage-processing task, respectively. Among the 202 remaining participants, 178 participants completed the reasoning test.

The descriptive statistics for the spans and the results of the reasoning matrix test can be found in Table 2.1. The spans were calculated with the partial unit scoring. This score is calculated as the sum across the trials of the proportion of correct recall in the correct serial position within trials, divided by the number of trials per list length. For instance, using three trials per list length, if a participant failed only one trial at length 3 (because the second item was not correctly recalled) and then completely failed at the remaining trials of greater lengths, then the participant would be granted a span of $2+(\frac{2}{3}\times1+\frac{1}{3}\times\frac{2}{3}) = 2.89$, with the first term 2 representing a span of 2 for the correct responses across list lengths 1 and 2, and $(\frac{2}{3} \times 1 + \frac{1}{3} \times \frac{2}{3})$ representing performance at length 3 (or: $(1{+}1{+}1)/3{+}(2{+}2{+}2)/6{+}(3{+}3{+}2)/9 = 2.89$). This scoring method allows to equate the weight of the different set sizes (see. Conway et al., 2005; Kane et al., 2004). The reasoning test scores were calculated as the sum of all the correct answers.

Table 2.1: Descriptive statistics for span tasks ($N = 202$) and fluid intelligence ($N = 178$).

	Mean	SD	Range	Skewness	Kurtosis
Processing-Storage	5.68	0.96	5	-1.12	1.16
Storage-Processing	5.45	0.94	5	-0.85	0.54
Matrix Reasoning	19.54	3.54	19	-0.87	0.52

We first analyzed the difference between the two types of tasks using both null hypothesis significance testing (NHST) and Bayesian analyses[3]. We provide the Bayes factors using equal priors for the considered alternative hypotheses to indicate their relative plausibility, knowing that a Bayes factors above 3 or below $\frac{1}{3}$) is substantial. Effectively, Bayes factors offer a simple expression of the degree to which data provide evidence for competing hypotheses across statistical tests (R. D. Morey, Romeijn, & Rouder, 2016).

The results showed that the mean span for the processing-storage task was found to be significantly higher compared to that of the storage-processing task ($t(201) = 4.46, p < .001$; $d = .31$; $95\%\ CI = [.17, .46]$). Unsurprisingly, the Bayesian paired t-test showed also evidence in favor of higher spans in the processing-storage condition ($BF_{10} = 911.3$).

Then, we analyzed performance as a function of item position, as we expected variations of performance due to the structural difference between the tasks caused by the processing-storage shift at the first and last positions. From this shift, we at least expected both higher performance at the first and last position for the processing-storage order, since the processing event could not disrupt the following storage event. However, the TBRS2 predicts a more complex pattern of differential performance from the first item to the last item, depending of the type of task. Figure 2.3 shows a gradient of systematic differences depending on item position, as the TBRS2 predicts that differences progressively increase until mid-list and then decrease until the last item. Our results effectively showed a gradient of systematic differences depending on item position, but the differences tended to progressively decrease until mid-list and then increased until the last item. To capture the interactions between item position and type of task, the data was

[3]While the orthodox NHST remains a common tool for drawing statistical inference from a sample, it has been widely criticized over the years mainly because of risks of type-I errors given the null hypothesis (e.g. Cohen, 1994). More recent approaches advocate for more appropriate methods developed by bayesian psychologists. The bayesian statistics can assess the relative plausibility of the null and alternative hypotheses while avoiding the several drawbacks of the NHST paradigm (Dienes, 2011; Gallistel, 2009; Wagenmakers et al., 2018). All of the Bayesian statistics were run in JASP (retrieved from http://jasp-stats.org/) with default parameters.

analyzed with a linear mixed-effects regression model in R using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). We did not aggregate the data, and thus we used the binomial family option to account for performance at each of the three trials per list length, depending on whether the recall of the item was correct or not at each position. For each sequence length (from 3 to 7 items), we compared five different models, depending on which factor was entered in the model (intercept alone, item position alone, type of task alone, item position + type of task, and item position * type of task). The participants were included as random intercepts in all of the models. Different increasingly complex models were tested by comparing their Akaike Information Criterion (AIC) values. The most parsimonious model (i.e. the best model with the lowest AIC) was chosen using the aictab function (library AICcmodvag). This procedure ranked as best the model including both factors and the interaction term for $L = 4$, $L = 6$, $L = 7$ with $p < .001$, and $L = 5$ with $p < .05$. The selected model for $L = 3$ was the one that included item position and type of task without their interaction.

Finally, we analyzed the relation between the two types of tasks and our measure of fluid intelligence. The correlations between span tasks and the reasoning subtest are presented in Figure 2.4. The span for the processing-storage condition correlated significantly with the reasoning test ($r = .230, p < .002$), as did the storage-processing task ($r = .259, p = 0.001$). The Bayesian correlations showed evidence for a correlation with intelligence higher than zero for the processing-storage condition ($BF_{10} = 10.73$) and for the storage-processing condition ($BF_{10} = 40.16$). The spans in both processing-storage and storage-processing tasks correlated significantly with each other ($r = .722; p < .001$). The Steiger's test for dependent correlations showed no significant difference between the correlation between the storage-processing complex span task and the reasoning test and the correlation between the processing-storage complex span task and the reasoning test ($r_{diff} = .05, z = 1.0255, p = 0.305$).

We also attempted to use an index of the relative difficulty of the tasks by selecting for each participant their average performance for the first item positions (i.e. how well they perform for the first item in the storage-processing minus the processing-storage condition). Again, we observed no significant correlation that could have indicated that the participants who were the most sensitive to the increased difficulty of the storage-processing complex span task tended to have a lower IQ.

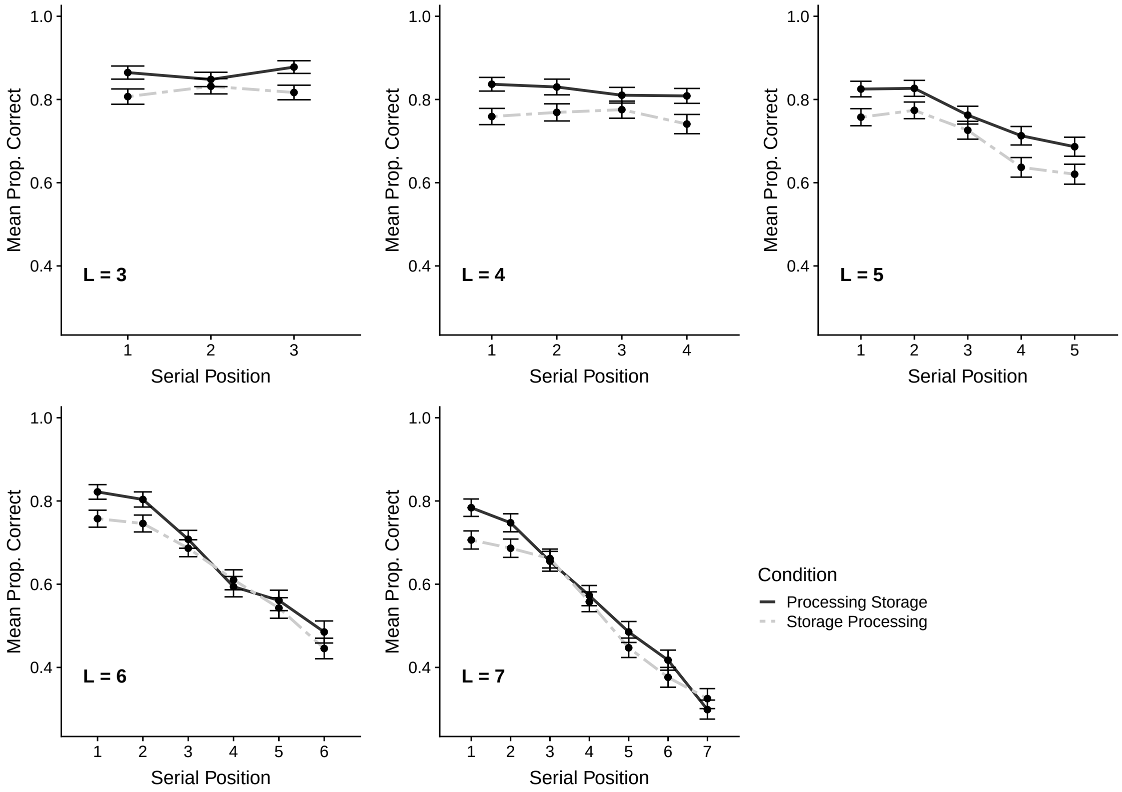


Figure 2.3: Serial position curves as a function of type of task.

To better approach the relation between the two types of complex span tasks, we conducted a 3-step mediating analysis by considering that the correlation between the Processing-Storage type of task and intelligence is mediated by the Storage-Processing type of task. The hypothesis was that this type of analysis could reveal an existing structural relation between the types of tasks. The regression coefficient between the Processing-Storage task and intelligence (direct effect) was significant $(.8430, p = .002)$. The indirect effect of the Processing-Storage task on intelligence was $(.72531) \times (.7092) = .51$, the value .72531 corresponding to the significant regression coefficient ($p < 2e - 16$) between the predicting Processing-Storage task and the resulting Storage-Processing task, and the value 0.7092 corresponding to the regression coefficient ($p = .07$) between the Storage-Processing task and intelligence in the multiple regression (i.e. controlling for the Processing-Storage task).

The significance of the indirect effect was tested using the R function mediation. Confidence intervals for the unstandardized indirect effects were computed based on 2000 bootstrap samples. The 95% confidence interval ranged from -0.00765 to 1.07 (the average bootstrapped unstandardized indirect effect was .52). This

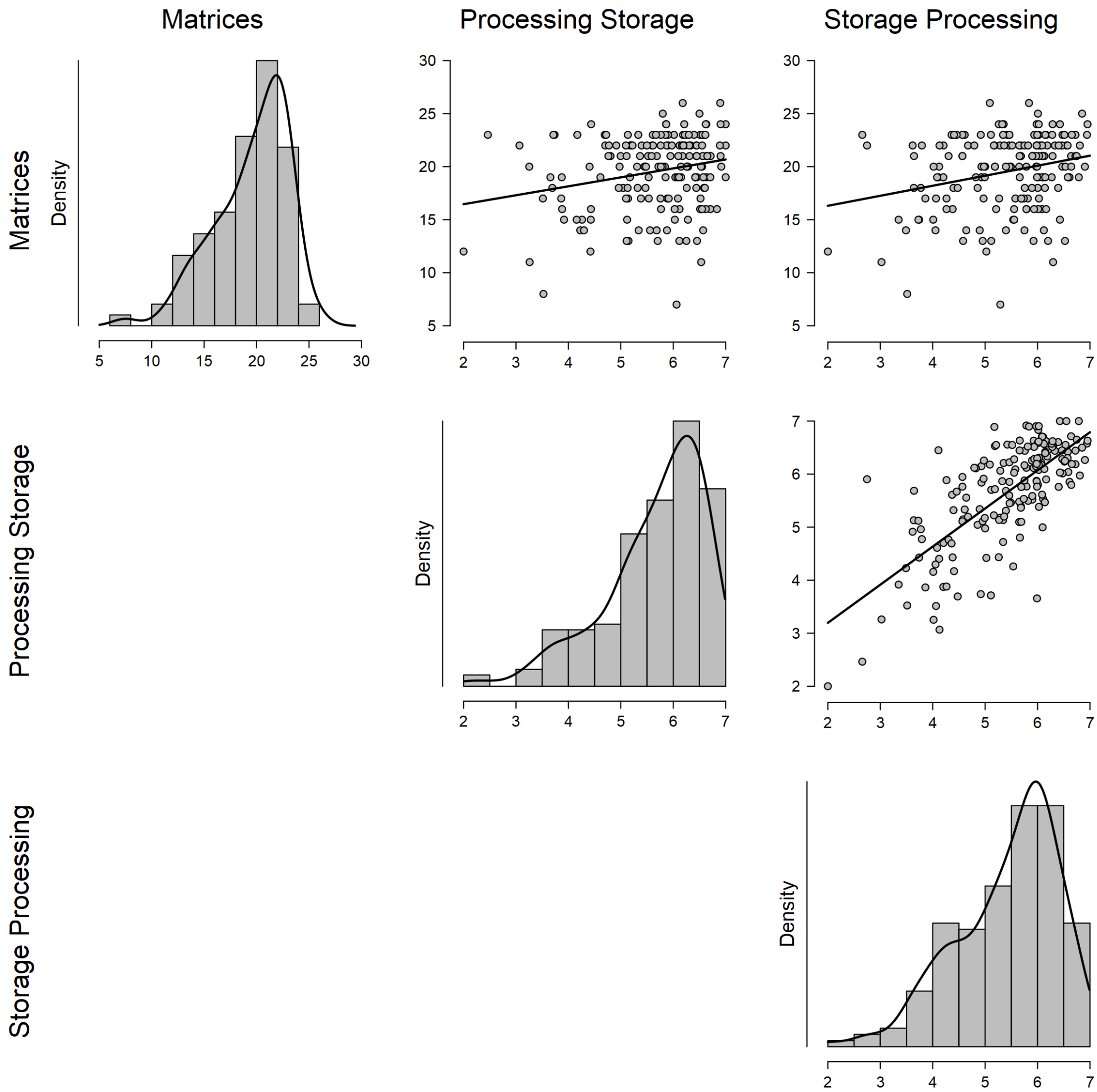


Figure 2.4: Bayesian correlation chart for the complex span tasks (Processing-Storage or Storage-Processing) and the matrix reasoning subtest ($N = 178$).

indirect effect was short of significance ($p = .05$) and the proportion mediated was equal to 61%, resulting from the values .52 (the indirect effect) divided by .84 (the total effect). In contrast, the average direct effect coefficient was equal to .32. This result means that the Processing-Storage type of task has a lower direct effect on intelligence.

2.2.3 Discussion

The current study examined whether the structure of complex span tasks (i.e. based on processing-storage vs storage-processing cycles) could influence the measure of WM capacity and whether these variants could impact the predictive power of higher-order abilities. Our hypothesis of a more difficult task in the storage-processing condition was derived from simulating the TBRS model. Our experiment confirmed our hypothesis that the structural pattern of the complex span task modulates working memory capacity, with greater spans being obtained when the task starts with a processing event instead of a storage event. The two conditions also influenced differential recall patterns as a function of item position, but not exactly the way predicted by the TBRS2.

One explanation for this discrepancy between the expected serial positions and the observed data might be due to the way the model handles refreshing. The original TBRS model does not make any assumption on the refreshing schedule nor on its duration. Gauvrit and Mathy (2018) discuss in their paper several variants of how refreshing could occur after an interruption. For the sake of clarity, we only presented in our introduction the prediction of the simplest variant (similar to the cumulative schedule of refreshing in forward order also implemented by Oberauer & Lewandowsky, 2011). In this variant, items are refreshed in their order of presentation, that is, always starting with the first item after an interruption. Other variants of the TBRS2 model can account for other refreshing schedules thanks to which the model is able to keep track of the last refreshed item or the least activated item but our tentative simulations did not lead to satisfactory results as the model in its present form does not enable to restrict the parameter search. However, theoretically, in a complex span task such as the operation span task, the ’last refreshed item’ variant would lead to a strong recency effect because refreshing would favor the last items of a list. In contrast, the ’least activated item’ variant would predict no clear serial position effects, as all items should be somewhat equally activated. As seen in Figure 2.3, except at very short list length, the data indicates a strong primacy effect and no recency effect. For this reason, it is unlikely that these variants could account any better for the data observed.

Another parameter of the TBRS2 concerns how long items are refreshed in working memory. Our predictions were based on a fixed refreshing duration (e.g. every item is refreshed for $0.08s.$, also implemented by Oberauer & Lewandowsky, 2011). However, the model has an alternative option, where refreshing goes on until a threshold of activation is met for a given item (or after a minimum duration if the threshold is already met for a hypothetical already highly activated item, for instance $0.01s.$). This variant leads to greater variability in the final product of recall compared to a fixed refreshing duration, especially if the number of items to remember is higher than the individual span. However, although such a variant is even more hazardous when simulated, we argue that it is not psychologically plausible, as it would require an extra component allowing the participant to constantly scan the activation level of the items to decide whether the refreshing process must be pursued for a given item.

We presented here four optional versions of the TBRS2, which appear unlikely to better account for the observed patterns of recall than the default version. However, many other modeling choices could have been made regarding the refreshing schedule, including its total duration or a limited number of items being refreshed at once (see Lemaire et al., 2018 for various computational implementations of refreshing schedules of the TBRS), but also the exact choices of the decay function. While it is likely that several other modeling choices combined could reproduce the data observed, again, finding the optimal set of parameters seems premature. It is also complicated to describe verbally exactly how these implementations would behave without conducting a thorough computational modeling study, as all these parameters interact with the design of the task such as its pace and list length (Lemaire et al., 2018). The fact that refreshing and decay processes are underspecified is an inherent limitation of verbal models in general. In contrast, computational modeling require that all parameters of a model are to be considered and well defined. This inevitably leads to a family of models that eventually predict different effects (Lewandowsky & Farrell, 2010), but discussing the predictions of every variant of the TBRS is beyond the scope of this dissertation.

Concerning the prediction of higher-order abilities, the structure of the complex span task did not influence the predictive power of the tasks. The mediation analysis helped partition direct and indirect effect of the types of tasks, but it remains difficult to interpret. Nevertheless, this result appears reassuring given all of the studies which have been conducted so far to estimate the relation between working memory capacity and general intelligence.

A potential concern with our correlational analysis is a ceiling effect, as the lists were limited to 7 items. However, only 5 participants obtained a span of 7 based on the partial unit scoring system. We reran the analysis using a more stringent scoring method (i.e. absolute scoring) and we did obtain reduced performance and reduced skewness, but the overall pattern of statistical results did not change. Therefore, similar conclusions could be drawn. Our result could be limited in that the higher-order abilities were assessed by a single subtest, and because the two experimental conditions were limited to 15 trials (following Unsworth et al., 2005). Further tests should attempt to generalize our findings by using several measures of fluid intelligence and using a greater number of trials per condition to reach a more precise estimate of the individual's spans. Additionally, such a project would ideally involve a larger number of complex span tasks using more diverse material and different types of concurrent tasks.

2.3 Conclusion

The results of our first experiment revealed several interesting findings. First, the structure of the task has an influence on the *estimate* of working memory capacity. When the processing episode is at the beginning of the task, the estimate of the span is increased. This finding is in line with the TBRS model prediction, as the first processing event does not have an impact on the cognitive load of the task. Our second result is that the prediction of intelligence was not affected by the structure of the task when the tasks were compared separately (i.e. the correlations between the types of tasks and intelligence were comparable). However, a mediation analysis can reveal a relation between the types of tasks as the Processing-Storage type of task appears to have a lower direct effect on intelligence than the Storage-Processing type, but this relation is not easy to interpret theoretically.

Regarding the main question of interest of the present dissertation(i.e. the effect of switch costs in working memory), both tasks are in theory suitable to explore our research question. However, because the storage-processing structure is more demanding, we will favor this type of task in the rest of this work. Moreover, it is worth noting that the need for a more strict control of the temporal aspects of working memory tasks has led the authors of the TBRS model to design simpler complex span tasks. The traditional complex span tasks widely used in the literature (e.g. reading span, operation span task) are self-paced in nature. As such, each processing step (e.g. operation to be solved) is displayed as long as the participant presses a key response that will in turn trigger the next step of the task. However, because the TBRS model supposes a constant rapid-switching between the two main functions of working memory, these self-paced tasks are unsatisfactory for the rigorous control of the temporal boundaries between storage and processing activities (Barrouillet & Camos, 2014b). Moreover, the discovery of the cognitive load effect demonstrated that complex span tasks do not need to be complex in order to highlight the existence of interference between storage and processing, they just need to capture attention. Therefore, tasks requiring time-constrained elementary activities, such as magnitude, judgment, reading aloud, parity judgment allowed to adequately measure complex span with the advantage that these activities rely less on academic skills compared to the traditional complex span task, and they happen to be just as predictive of higher order abilities (Lucidi et al., 2016).

Following this line of reasoning, we will pursue in the next chapter our investigation of the serial structure of the complex span task at a finer level with the use of elementary processing steps presented at a predefined pace to manipulate switching.

3

Variations of processing schedules in complex span tasks

The complex span tasks used in working memory studies typically involve the regular alternation between storage and processing activities. The storage task is the main task, as the measure of interest in the complex span task is the capacity of individuals. Capacity is generally estimated by the number of correctly recalled items in the correct order. The processing component of the task is triggered by a concurrent task that consists in processing distractor items in between the to-be-recalled stimulus items.

Crucially for the present thesis, the structure of a processing activity in a complex span task is –generally– strictly regular. A typical concurrent task structure could be described as follows : 10101010 in which 1 stands for a distractor and 0 for a constant interstimulus interval. In other words, in the vast majority of the complex span tasks used in the literature, attention is directed away at fixed intervals. However, this task design may introduce a confound between the processing efficiency of participants and their switching efficiency. The first one refers to the capacity to process items rapidly in order to resume to the main task to avoid any further memory loss, while the second one relates to the capacity to alternate rapidly between processing and storage, independently of their speed of processing. This confound could unwisely lead to grant special privilege to the role of attentional refreshing in comparison to executive functioning.

Theoretically, a detrimental effect of switching on memory performance could affect the construct validity of the complex span task, and impact not only the theory but also interpretation of data. For instance, regarding developmental data, because children have not fully developed executive functions, this factor could be a confound to interpret their lower performance in complex span tasks. The rational here, is that older participants could perhaps perform better at regular tasks because they are able to better anticipate the schedule of the processing steps in order to refresh their memory traces at appropriate times. Previous work has provided credible evidence supporting a detrimental effect of task switching on working memory performance (Liefooghe et al., 2008). However, switch costs in working memory have been investigated so far by manipulating the amount of switches between two processing tasks. Therefore, it remains unclear whether switching between storage and processing (and the other way around) has also a detrimental effect on working memory.

The aim of the following studies was to explore switching costs in complex span tasks consisting of pseudo-randomly organized processing events along with trials of different set sizes. The idea was to test patterns of distractors which are unpredictable for the participant in order to examine if working memory performance can be impaired by switching costs between storage and processing (independently of the duration of the attentional capture associated to the execution of the concurrent task). Recall that we consider that switches are determined solely by the structure of the task (represented in Figure 1.17a) . In other words, each time the concurrent task displays consecutively a distractor and a blank delay (or the other way around), we posit that a switch between storage and processing activities occurs. We expect less accurate performance on both the recall and the concurrent task when the number of switches to execute is high compared to trials in which the events are grouped together.

The alternative hypothesis is based on the TBRS model, performance should only be related to the cognitive load of the task at hand. More precisely, switch costs between storage and processing activities may only be, at best, tacitly assumed to to be part of the attentional capture due to the processing task, but they cannot be conspicuously demonstrated by varying the structure of the task.

3.1 Experiment 2

3.1.1 Method

Participants. Seventy young adults of Université Fanche-Comté (mean age = 20.70 years, $SD = 2.07$) participated to this experiment in exchange for course credit. The duration of the experiment was approximatively one hour. All participants were tested in a quiet room at the university. This unpublished data was collected by Fabien Mathy.

Material and Procedure. The complex span task used in this experiment consisted of ten trials of 2 to 6 letters randomly chosen without replacement from the following set of consonants: B, F, H, J, K, L, P, Q, R, S, V, X. Each trial started with a fixation cross displayed for 2000 ms centered on the screen followed by the letter to be remembered presented for 1000 ms. After each letter, a concurrent task took place in which the participant was invited to press on the space bar whenever the digit 1, 2 or 3 was presented on the screen. These processing events were randomly drawn from a pool of digits from 1 to 9. The total duration between two letters was randomly drawn between 1 s and 5 s (this duration could vary between letters within trials to make the task the least predictable). During these intervals, digits were displayed for either 1000 ms or 500 ms depending on the condition of the participant (the chosen condition was applied for the whole session for a given participant[1]).

For example, if an interval of 3 s was randomly drawn after one letter and if a duration of 1000 ms was chosen for the display of one digit, the following patterns of events could take place: 111, 000, 100, 001, 010, 011, 101, 110, where 1 stands for a period during which a digit could be displayed, and 0 stands for a distractor-free time. After each letter the participant could be administered a new pattern. For instance, in a set size of three letters, the participant could see the following global pattern: L0011L001L11111 (where L stands for a letter), as seen in Figure 3.1. For this pattern, the duration after the successive letters was 4 s, 3 s, and 5 s, respectively. However, it is worth noting that the likelihood of the presentation of digits was determined by the cognitive load also randomly drawn as explained below.

The cognitive load of a given trial was constrained to vary from easy to difficult based on six different theoretical probabilities of randomly displaying a digit between two letters (i.e. p_d = 0, .20, .40, .60, .80, or 1). Depending on this probability, the possible patterns did not have the same chance of being drawn. For example, if a concurrent task of 3 s was randomly drawn after one letter and if a duration of 1000 ms was chosen for the display of one digit, and if $p_d = 0$, the only pattern that could be built was 000, among the set 111, 000, 100, 001, 010, 011, 101, 110. The chosen probability was applied to all events of a given trial. For instance, given a probability of .20, each slot of 1000 ms or 500 ms time could be filled with

[1]Note that because the 1000 ms *vs.* 500 ms factor was not intended to interact with our main factors, we did not intend to analyse its effect. The only goal of this manipulation was to avoid the risk that the task would be too slow or too rapid. This factor was not significant in the mixed models we run in the main analysis on memory performance whatsoever.

a digit with a probability of .20. The equal cognitive load was thus expected for the entire sequence, that is, a cognitive load of .20 was expected established of individual probabilities of .20 based on the binomial distribution. In sum, the essence of the task was to make the events appear for the participant as random as possible, after the cognitive load was set for a given trial. The cognitive load was chosen before randomly drawing the events to balance the levels of cognitive load uniformly (i.e. p_d = 0, .20, .40, .60, .80, or 1). If we had not used this procedure, the binomial distribution of the events around $p = .50$ would have produced a normal distribution around .50, and in this case the procedure would have generated the cases such as p_d = 0 or p_d = 1 too rarely.

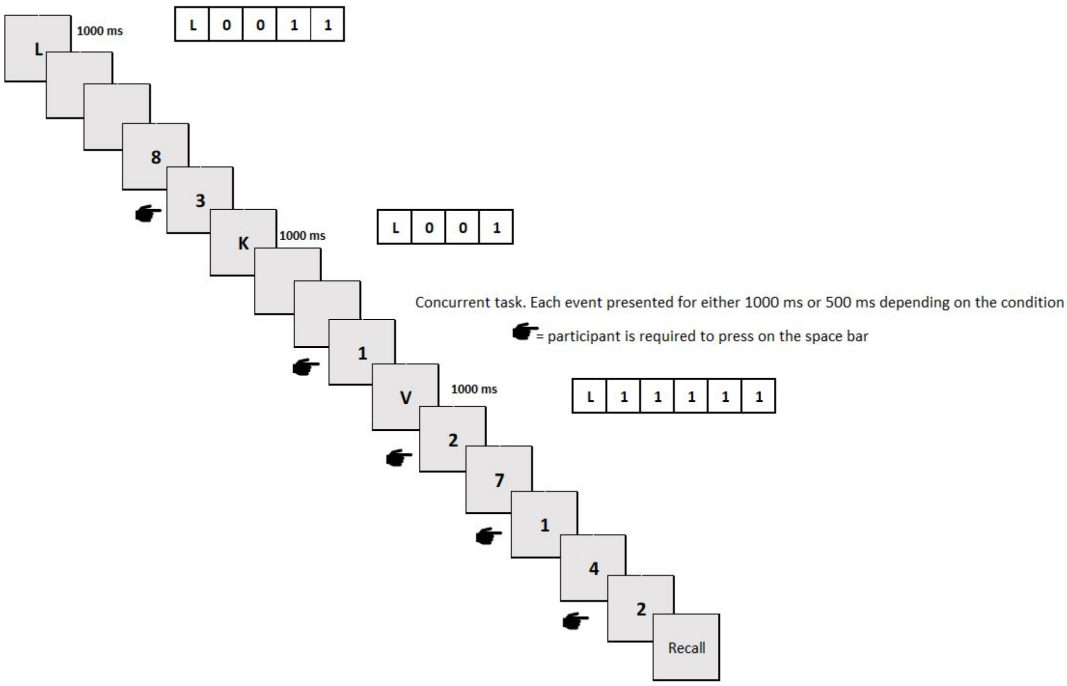


Figure 3.1: Example of variations of the structure of the processing events for three letters to be recalled in Experiment 2, and the corresponding notation of these patterns used throughout the present dissertation. L stands for the presentation of a letter, 1 stands for a distractor (i.e. digit), and 0 stands for a free delay.

After each trial, the participants were asked to recall all of the letters presented in the correct order, if possible. The responses were given by the participants by clicking on a visual keyboard 3×4 letters. The participants were not able to correct their answer once they clicked on the letter (to avoid output interference). A feedback was provided on the next screen, and the word GO was finally indicated to let the participant know that they could move on to the next trial by pressing the space bar again. Before the beginning of the complex span task, the participants were invited to practice on 18 trials with progressive difficulty (i.e. increase of the cognitive load). During the training trials, feedback on the successful performance of the concurrent task was given by a visual cue (i.e. a green flash indicating correct key press, and a red flash whenever a commission error was made).

3.1.2 Statistical analysis method

For this experiment and the following ones, model comparison analyses were based partly on the Akaike information criterion (AIC, Akaike, Petrov, & Csaki, 1973), which is an estimator of the quality of a statistical model. Simply put, the AIC criterion evaluates the quality of the fit of the model and its simplicity. In our analysis, we will use the AICc criterion, which is a correction of the AIC for smaller sample sizes (usually $N < 40$). When sample sizes are more important the AIC and the AICc are similar because the AICc converges towards the AIC. For consistency reasons throughout our different experiments we will use only the AICc criterion, even when the sample size can be considered large.

Additionally, we will use the weight of a model to determine if a model is better than another one. More specifically, from the weight of a model we will be able to calculate the evidence ratios between two models, which is the ratio between the weight of two different models. The weight of a model is always within a range of $[0; 1]$, thus evidence ratios will have a range of $[0; \infty[$. A high evidence ratio in favor of one model indicates that it is in all likelihood the best model (i.e. the more parsimonious) to account for the data. We will use as guideline the grade range of evidence ratios established by Burnham and Anderson (2002). Table 3.1 summarizes these guidelines.

Table 3.1: Summary of grades of evidence for two models M_1 and M_2.

Grades of evidence	Evidence Ratio
Equal Support	$1 - 2.7$
Support for M_2, but do not discard M_1 yet	$2.7 - 20$
Very strong evidence for M_2. M_1 can be rejected	> 150

3.1.3 Estimates of statistical fit based on AIC values

The AIC criterion is based on information theory, according to which a model describing a given set of data will almost always lack precision. The AIC evaluates this loss of information based on a basic rule: the fewer information is lost, the better the quality of the model in question.

For instance, if the two AIC values are 100 and 102 for two respective models M_1 and M_2, then we consider that it is $exp((102 - 100)/2) = 2.7$ more probable that the first model minimizes the loss of information. Because the AIC value depends on both the negative value of the log of the likelihood and the number of parameters, if the likelihood is high, for instance .0001, then $-log(.0001) = 9.2$ is low. If the model has 4 parameters, then the final AIC value is penalized by the complexity of the model and $AIC = 2 \times 4 + 9.2 = 17.2$. If the likelihood of another model is lower, for instance .00001, then $-log(.00001) = 11.5$ is higher, but if the model has only 2 parameters, the final AIC value is less penalized by the complexity of the model and $AIC = 2 \times 2 + 11.5 = 15.5$. In that case, the most satisfying fit is offered by the simplest model (see Pitt & Myung, 2002) which then is considered the most parsimonious model.

The best model here is a compromise because there could always be a highly complex model accounting for all of the data up to a point where the noise in the data would also be explained by the model, which is not what is expected by a model. Rather, we are looking for a model that captures the structure that generated the data, not some noise in the data.

3.1.4 Results

To determine if switching between storage and processing activities during a complex span task might be more time-consuming than originally thought by the TBRS model, we conducted separate analysis examining the effect of switching on recall and concurrent task performance. In both analyses, results from one participant was excluded as the proportion of correctly recalled items and performance in the concurrent task deviated more than two standard deviations from other participants. For the remaining 69 participants (mean age = 20.7 years, $SD = 2.08$) the average success rate on the concurrent task was 75% ($SD = 0.17$).

Analysis of recall performance. We used the proportion of correctly recalled letters in the correct serial position within a trial as our scoring method for the memory task. This score enabled us to create our main dependent variable referred to as memory performance from now on.

We analyzed memory performance as a function of the independent variables in our experiment (i.e. memory load, cognitive load, and the mean number of switches between two letters). We used the *lme4* (Bates et al., 2015) package in R to implement linear mixed models, in which the independent variables were considered as fixed effects. The participants were entered in the models as random effects in order to account for their individual differences and the fact that repeated measures were taken from each participant.

In total, we analyzed four different models to test our hypothesis according to which switching from storage to processing during a complex span task hinders memory performance. To this end, we compared two simple statistical models that included only memory load (i.e. list length of the letters to be recalled) and cognitive load as fixed factors (with and without the interaction term, referred to as $M1$ and $M2$ respectively), and two enhanced models that included additionally the factor mean number of switches between two letters (with and without the interaction term, referred to as $M3$ and $M4$ respectively). A summary of our four statistical models is presented in Table 3.2.

Table 3.2: Summary of the statistical models created for this study

$M1 \longleftarrow$ DV$\sim$ List Length + Cognitive Load + $(1|Subject)$
$M2 \longleftarrow$ DV$\sim$ List Length * Cognitive Load + $(1|Subject)$
$M3 \longleftarrow$ DV$\sim$ List Length + Cognitive Load + Mean number of switches + $(1|Subject)$
$M4 \longleftarrow$ DV$\sim$ List Length * Cognitive Load * Mean number of switches + $(1|Subject)$

Table 3.3 shows a ranking of the models in terms of best fit, and the details of the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. These values were calculated with the function AIC*tab* of the AIC*cmodavg* in R.

The statistical analysis indicated that the model incorporating length list, cognitive load and the mean number of switches between two letters and their interactions was the most likely model given the observations. However, before detailing further the parameters of interest of this model, it is worth noting that there is only a small difference between the AICc of the first two models, as seen from Table 3.3. The calculated evidence ratios between these models indicated equal support for both models ($M4/M2 = 1.34$). Therefore, the factor mean number of switches added to the analysis is not significantly beneficial.

Table 3.3: Summary of model selection based on the number of parameters K, AICc values, delta AICc values and weight of each model. The dependent variable is memory performance.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M4$	10	-672.97	0	0.57
$M2$	6	-672.39	0.58	0.43
$M1$	5	-654.82	18.15	0
$M3$	6	-652.86	20.11	0

*Note: M4 = List length * Cognitive load * Mean number of switches ;M2 = List Length * Cognitive load ;M1 = List length + Cognitive load ;M3 = List length + Cognitive load + Mean number of switches*

Regarding the parameters of interest of the model best ranked, we unsurprisingly observed an effect of memory load, with higher performance on the recall task for shorter list lengths compared to longer list lengths ($F(1, 4073.2) = 46.71$, $p < .001$). The estimate showed a -3% memory performance decrease when memory load was increased by 1 stimulus item. Additionally, in line with the TBRS model, the cognitive load of the task also influenced memory performance, with lower performance on higher CL conditions compared to lower CL ($F(1, 4077.7) = 4.25$, $p < .05$). The estimate showed a -9% memory performance decrease when cognitive load was increased by 100%. Furthermore, a significant interaction between memory load and cognitive load was found ($F(1, 4083.8) = 26.93$, $p < .001$). Although, robust effects of list length and cognitive load on working memory capacity have already been established by previous researches (Barrouillet et al., 2004, 2007; Barrouillet & Camos, 2012), our result confirms the generalization of these effects when the structure of the task is unpredictable for the participant. Finally, against our hypothesis, recall performance did not differ significantly as a function of the number of switches to perform between storage and processing activities ($F(1, 4079.7) = 2.29$, $p = .13$).

The overall interaction between memory load, cognitive load and the mean number of switches to perform was significant ($F(1, 4078.9) = 6.76$, $p < .01$), as was the interaction between the mean number of switches and cognitive load ($F(1, 4079.6) = 4.54$, $p < .05$). Figure 3.2 shows memory performance as a function of the mean number of switches between two letters, split by cognitive load (categorized as quartiles for more readability). As seen, memory performance tends to slightly decrease as both the number of switches and cognitive load increases.

However, it is worth noting that the distribution of memory performance was negatively skewed (79% of the trials were perfectly recalled by the participants), indicating that the task was most likely too easy. Therefore, it is possible that the effect of switching was not apparent on its own because of this ceiling effect. Lastly, no significant interaction was found between the mean number of switches to perform and memory load ($F(1, 4077) = 3.43$, $p = .06$).

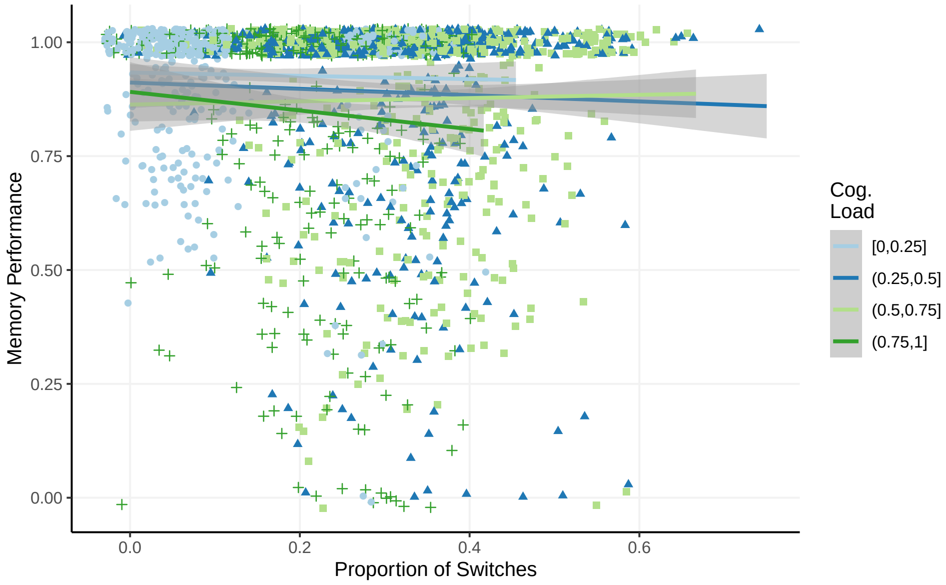


Figure 3.2: Mean memory performance as a function of cognitive load and proportion of switches between two letters in Experiment 2

As shown in Figure 3.3, the relation between the mean number of switches between two letters and cognitive load is described by an inverted U-shaped curve. In extreme cognitive load conditions (low and high CL, respectively [0-0.25] and (0.75-1]), the structure of the task allows only few switching opportunities between storage and processing activities by construction. As seen in Figure 3.3, there is a confound between the mean number of switches between letters and cognitive load, across the four cognitive load quartiles but also within each of the quartiles (each of the four regressions turned out significant, but the two regressions for (0.25-0.50] and (0.50-0.75] were the smallest). Hence, the most conducive condition to explore the effect of switching is a medium cognitive load condition (between .25 and .75), in which the confound is the smallest. For this reason, we examined further the relation between the mean number of switches and memory load on performance with a mixed model analysis including only trials with a cognitive load between .25 and .75. The results of this analysis showed unsurprisingly an effect of list length on memory performance $F(1, 4075.0) = 76.74$, $p < .001$, but no global effect of the mean number of switches ($F(1, 4082.2) = 2.56$, $p = .11$). Additionally, a significant

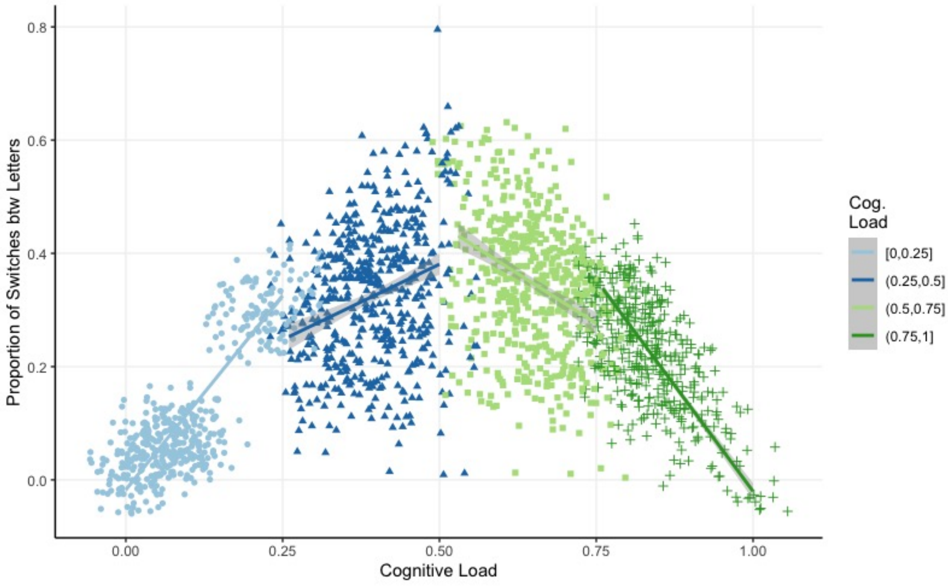


Figure 3.3: Proportion of switches between two letters as a function of cognitive load in Experiment 2

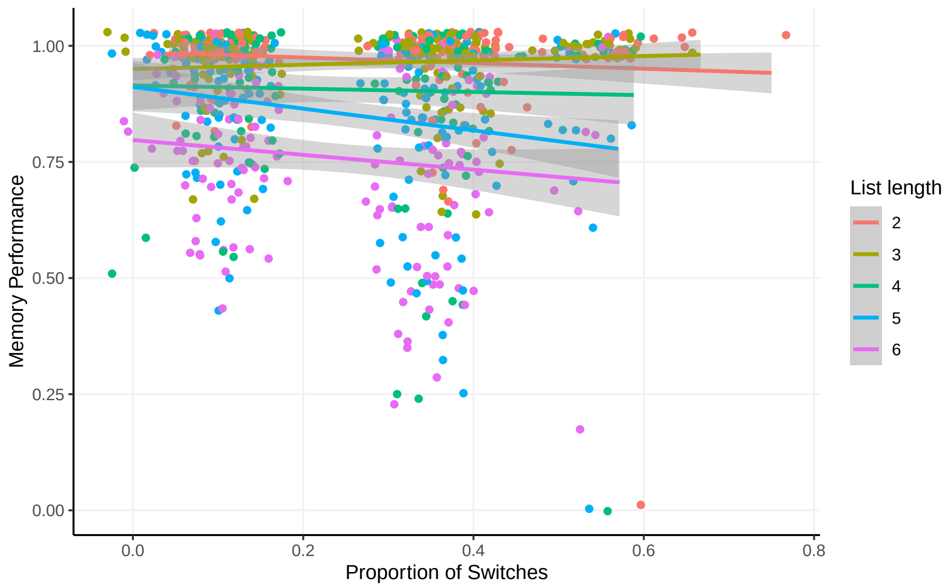


Figure 3.4: Mean memory performance as a function of the proportion of switches between two letters and memory load in Experiment 2, when only the trials of medium CL (0.25-0.75] were included.

interaction was found between list length and the mean number of switches on memory performance ($F(1, 4076.3) = 17.68$, $p < .001$). The significant interaction reflected a greater effect of the mean number of switches as list length increased, as seen in Figure 3.4. This effect of the mean number of switches was significant for all list lengths except for the length of three letters to be recalled.

Analysis of concurrent task performance. An alternative way to examine the effect of switching in our complex span task is to explore how the concurrent task was succeeded when the mean number of switches between two letters increases. The idea is that our manipulation could have impacted the concurrent task instead of the main task.

We used the proportion of correctly executed decisions in the concurrent task within a trial as our dependent variable. This dependent variable will be referred as sensitivity to the concurrent task from now on. Sensitivity was calculated by subtracting the proportion of false alarms from the proportion of hits within a trial. A false alarm corresponded to the situation in which the participant wrongly pressed the space bar when a decoy was presented (i.e. a digit between 4 and 9). The proportion of hits corresponded to the situation in which the participant correctly pressed the space bar when a target was displayed (i.e. a digit between 1 and 3). Sensitivity can theoretically range between the inverval $[-1; 1]$, a negative sensitivity means that the participant made more false alarms than target hits, whereas a sensitivity of 0 means that the participant omitted all targets and made no false alarms.

We proceeded to model comparisons as in our previous analysis on memory performance by analyzing sensitivity as a function of the independent variables in our experiment. The structure of the statistical models were identical to our previous analysis on memory performance (see Table 3.2 for a summary of the statistical models).

Table 3.4 shows a ranking of the models in terms of the best fit, and the details of the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. The statistical analysis indicated that the model incorporating length list, cognitive load and the mean number of switches between two letters without their interaction was the most likely model given the observations. Once again, there was only a small difference between the AICc of the first two models, as seen from Table 3.4. The calculated evidence ratios between the first models indicated support for the full model without the interaction term, although the interaction model is not discarded yet ($M3/M4 = 3.02$). Additionally, comparison between $M3$ and $M1$ indicated strong evidence in favor of the model incorporating the mean number of switches ($ER > 150$), which informs that this factor is an important when considering performance at the secondary task.

Table 3.4: Summary of model selection in Experiment 2 based on the number of parameters in each model (K), AICc values, delta AICc values and weight of each model. The dependent variable is the sensibility to the concurrent task.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M3$	6	882.25	0	0.75
$M4$	10	884.45	2.21	0.25
$M2$	6	912.72	30.48	0
$M1$	6	919.90	37.66	0

*Note: M3 = List length + Cognitive load + Mean number of switches ;M4 = List length * Cognitive load * Mean number of switches ;M2 = List Length * Cognitive load ;M1 = List length + Cognitive load*

Regarding the parameters of interest of the model best ranked, we observed a significant effect of all our factors of interest on sensitivity (list length $F(1, 2957.9) = 5.02$, $p < .05$; cognitive load $F(1, 2970) = 19.79$, $p < .001$; number of switches $F(1, 2636.7) = 43.81$, $p < .001$). The estimate provided by the mixed model was negative for list length.

From our previous analysis, we know that cognitive load and the number of switches are correlated to each other. Therefore, we examined further the relation between concurrent task performance and the number of switches and memory load when cognitive load was constrained between .25 and .75. This analysis revealed considerably different results: the effect of the number of switches on concurrent task performance was no longer significant.

3.1.5 Discussion

In order to test the possible confound between processing efficiency and switching ability, the present experiment used a complex span task in which the events of the processing task were not predictable. We posited that a greater number of switches between processing and storage activities would reduce performance on both recall and the concurrent task.

The best model is a compromise between the loss of information and the number of parameters of a model. Hence, between two models receiving equal support, the most parsimonious model is to be favored. Aside the basic effect of list length (which is a factor that is not disputed between theories and which is in essence used to measure the span), the results confirmed the robustness of the cognitive load factor. This result is novel because the task structure was subject to variations (i.e. the distractors could occur randomly in the concurrent task, and the duration of the concurrent task could not be predicted between two letters).

Regarding our main question of interest, our results showed no credible evidence in support the idea that the structure of a task can produce a detrimental effect of switching between storage and processing on complex span task performance. However, a closer look at the relation between cognitive load and the proportion of switches within trials revealed only few switching opportunities between storage and processing activities in the highest cognitive load conditions. When switching was examined in more conducive conditions (i.e. medium cognitive load), we observed, in favor of our hypothesis, a negative trend induced by switching for list lengths 2, 4, 5, and 6. This is a promising result, which suggest effectiveness of our manipulation of the switches. Our analysis of the concurrent task performance showed no effect of switching when the highest cognitive load conditions were removed from the analysis.

In conclusion, the present study attempted to randomize the complex span task maximally. Our findings indicated, however, that switching between storage and processing is best manipulable when cognitive load is moderate. The following experiments will pursue our efforts to detect switching costs between the two main functions of working memory while restricting the task at hand to this cognitive load range.

3.2 Experiment 3

The main goal of the present study was to detect unambiguous switch costs between storage and processing activities that are prompted by the structure of a complex span task. Our previous experiment failed to detect such costs. We attributed our lack of positive findings to two main causes. Firstly, the structure of the task of our previous experiment allowed only few switching opportunities between storage and processing activities by construction. Secondly, the task was most likely too easy to allow switch costs to be noticeable on performance. In light of our previous findings, we slightly increased the difficulty of the concurrent task by implementing a greater number of targets while restricting the range of the cognitive load to moderate values.

In order to maintain a constant cognitive load while manipulating the number of switches in a complex span task, we created six different processing schedules that vary on how distractors and interstimulus intervals are distributed within a trial. These patterns are shown in Figure 3.5. Akin to our previous experiment, we assumed that the processing schedules that involved numerous alternations between processing and storage events would induce switching costs by potentially impairing recall and concurrent task performance (pattern 3 and 6 in Figure 3.5).

Additionally, we also identified that some patterns might be considered as harder to learn because they did not present a fixed number of distractors in a row and fixed interstimulus intervals. Hence, we expected these irregular patterns to reduce both recall and concurrent task performance by rendering task-switch preparation more difficult (pattern 1 and 4 in Figure 3.5).

Finally, a closer look at the processing patterns invited us to assert one more prediction regarding a beneficial consolidation effect in working memory, although this was not explicitly part of the initial goal of our study. Consolidation is thought to stabilize and strengthen novel information (Jolicœur & Dell'Acqua, 1998). Numerous recent studies have put forward the idea that opportunities of undisturbed intervals of free time displayed immediately after the presentation of storage events should enhance post-encoding processes (e.g. Bayliss et al., 2015; De Schrijver & Barrouillet, 2017; Nieuwenstein & Wyble, 2014; Ricker & Hardman, 2017). In essence, this should translate to improved working memory when consolidation opportunities are provided. It turned out, that in our experimental material, consolidation processes could involve three patterns that start with a delay of free time immediately after the presentation of a storage item (pattern 1, 2 and 3 in Figure 3.5). In line with previous work, we expected to find better recall performance when a consolidation opportunity was provided. We also expected that consolidation would have a beneficial effect on concurrent task performance, since it is also known that responses are more likely to be slower and less accurate when consolidation is interrupted (Bayliss et al., 2015; Jolicœur & Dell'Acqua, 1998).

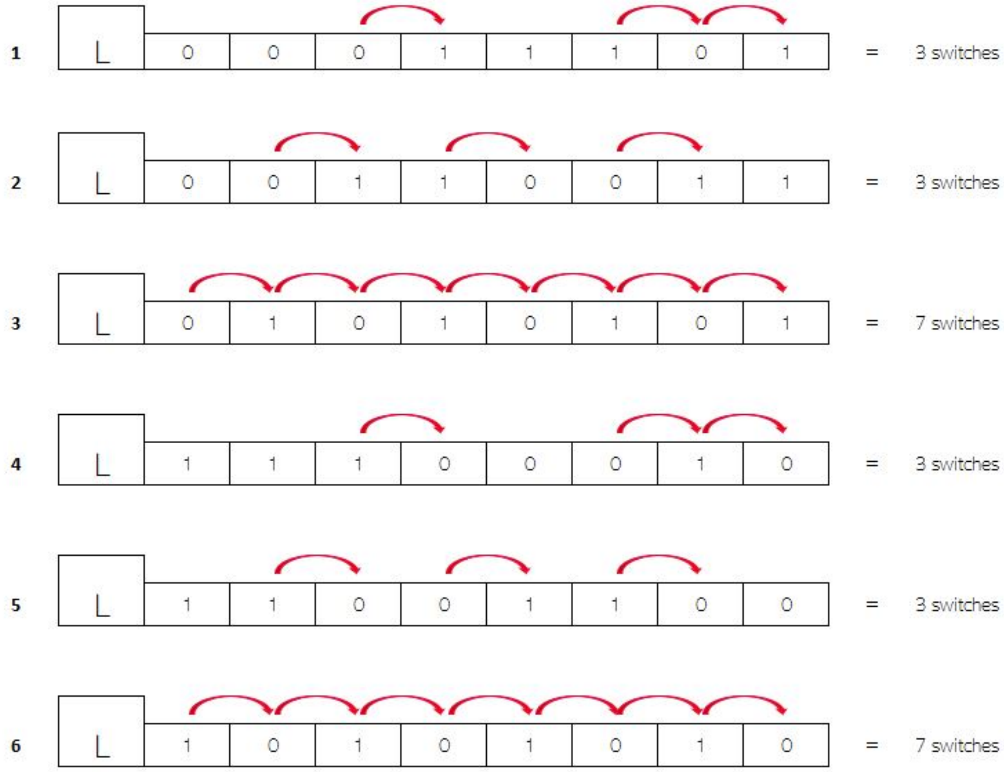


Figure 3.5: Processing schedules of the concurrent task used for Experiment 3. *Note:* L = Letter displayed for 1500 ms. Every processing event (i.e. free delay and digit) is presented for 800 ms.

3.2.1 Method

Participants. Sixty students (51 females, 9 males; mean age = 22.05 years, $SD = 4.15$) from Université Côte d'Azur participated to this experiment, in exchange for partial course credit. The duration of the experiment was approximately 50 minutes. All participants were tested in a quite room at the university. This unpublished data was collected by Cauchi (2016).

Material and procedure. Participants were invited to memorize in the correct serial order letters while completing a continuous performance task as in Experiment 2. The storage stimuli were identical to the previous study. Processing items were digits from $0 - 9$ drawn in a random order. The participants were instructed to press on the space bar when a digit between 5 to 9 appeared on the screen. In comparison to the previous experiment that only requested a press for 3 digits, we thought that increasing the number of presses would increase the attention demand in the concurrent task.

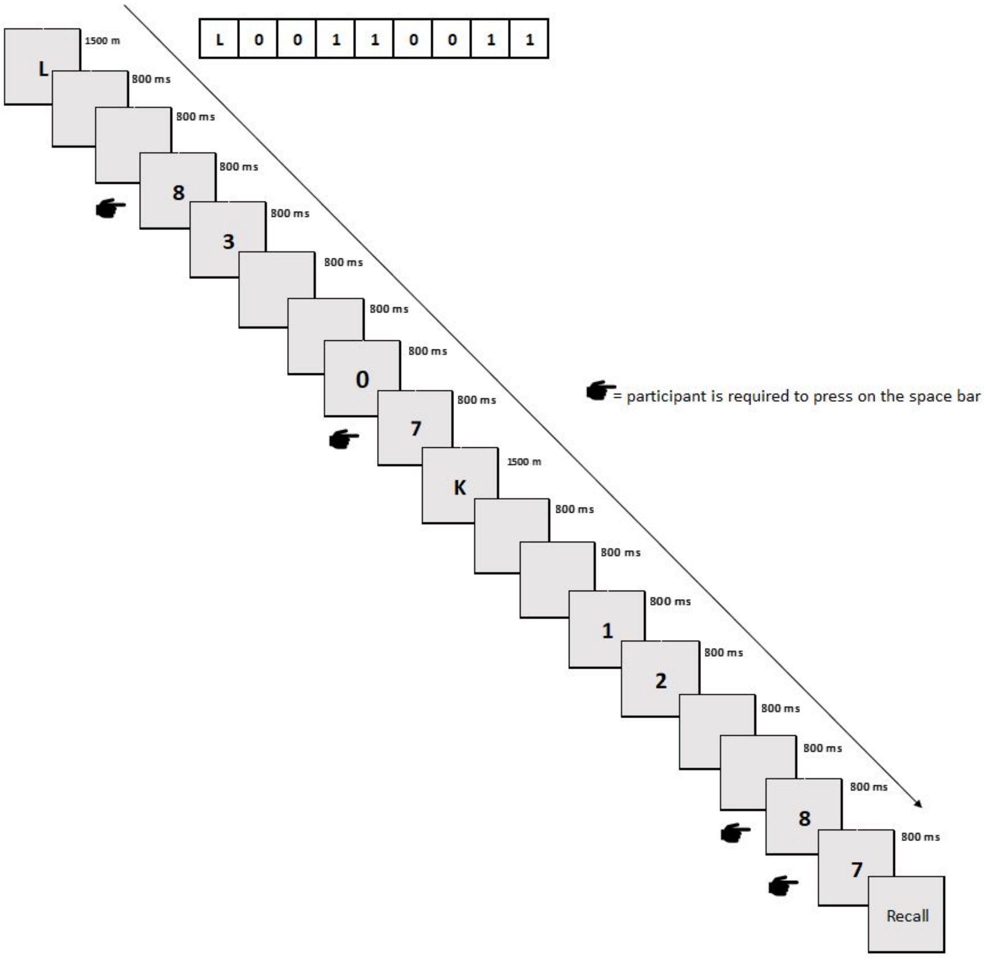


Figure 3.6: Example of variations of the structure of the processing events for two letters to be recalled in Experiment 3. The schedule of the processing events varied according to six predefined patterns. *Note:* L = Letter displayed for 1500 ms. 0 refers to a free delay and 1 corresponds to a digit, both displayed for 800 ms.

In contrast to Experiment 2, the patterns of the processing events were defined beforehand by limiting the range of the cognitive load and by fixing the interstimulus interval. Six different patterns were used, consisting always of four digits randomly drawn from 0 to 9 and four slots of free time. Thus, the cognitive load of the patterns was constant across trials. The main manipulation in the task was the sequential combination of digits and free time, which was determined by the pattern used. For each set size each pattern was repeated twice. Each pattern was randomly drawn without replacement at the beginning of a trial.

The length of the trials was progressive from 3 to 5 letters, with 12 repetitions per set size, making thus 36 trials in total. Each trial started with a fixation cross displayed for 750 ms, after which the first letter appeared for 1500 ms. After the letter, the processing events (i.e. four digit and four slot of free time) were displayed, each for 800 ms (see Figure 3.6). At the end of each trial, a matrix with all the possible consonants which could be presented during this experiment was displayed at the center of the computer screen. The participants were then invited to recall the letters in their correct serial order by clicking on a box next to the letter. Once the participant clicked on the letter, corrections were impossible. Before starting the experiment, the participants practiced for about 20 minutes. A feedback identical to Experiment 2 was provided during the training trials.

3.2.2 Results

We conducted separate analyses examining the effect of the processing patterns on the recall task and on the concurrent task, as in Experiment 2. Results from one participant was excluded from both analysis as the proportion of correctly recalled items and performance on the secondary task deviated more than two standard deviations from other participants. For the remaining 59 participants: mean age = 21.4 years ($SD = 2.4$).

Analysis of recall performance. We used the same scoring method as in the previous experiment, that is the proportion of correctly recalled items in their correct serial position. We will again refer to this dependent variable as memory performance.

We first conducted a linear mixed model analysis to see if the processing patterns had an overall effect on memory performance. We created five different statistical models. The null model ($M0$) only contained participants as a random factor. Then, two statistical models (referred to as $M1$ and $M2$) additionally included memory load (i.e. list length of the letters to be recalled) and processing pattern of the concurrent task as fixed factors. Finally, two more enhanced models included both fixed factors (with and without the interaction term, referred to as $M3$ and $M4$ respectively). A summary of our five statistical models is presented in Table 3.5.

Table 3.6 shows a ranking of the models based on their fit, the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. The results showed that the model incorporating only list length ($M1$) was best ranked, and thus considered the most likely model given the observations. The evidence ratio between this model and the null model was strong ($M1/M0 = > 150$).

Table 3.5: Summary of the statistical models of the main task in Experiment 3.

$M0 \longleftarrow$ DV$\sim$ + $(1|Subject)$
$M1 \longleftarrow$ DV$\sim$ List Length + $(1|Subject)$
$M2 \longleftarrow$ DV$\sim$ Processing pattern + $(1|Subject)$
$M3 \longleftarrow$ DV$\sim$ List Length + Processing pattern + $(1|Subject)$
$M4 \longleftarrow$ DV$\sim$ List Length * Processing pattern + $(1|Subject)$

Table 3.6: Summary of model selection in Experiment 3 for memory performance based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 3.5.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Data including all processing patterns				
$M1$	4	-1508.81	0.00	0.89
$M3$	9	-1504.67	4.14	0.11
$M4$	14	-1497.39	11.42	0.00
$M0$	3	-1424.48	84.33	0.00
$M2$	8	-1420.12	88.69	0.00
Patterns 00110011 *vs.* 11001100 only				
$M3$	5	-501.15	0.00	0.45
$M4$	6	-500.89	0.26	0.39
$M1$	4	-499.15	2.00	0.16
$M2$	4	-477.19	23.96	0.00
$M0$	3	-475.38	25.77	0.00

Against our hypothesis, the model including only the factor processing pattern was ranked last, which indicates that the explanatory power of this variable was null. Regarding the parameter of interest of the model best ranked, our results showed an effect of memory load, with higher performance on the recall task for shorter list lengths compared to longer list length ($F(1, 2058.3) = 88.2$, $p < .001$). The estimate showed a -4% decrease in memory performance when memory load was increased by 1.

To examine specifically switching effects on memory performance, we compared the processing patterns manipulating only the switch factor (i.e.11001100 *vs.* 10101010, including respectively three and seven switches). We purposefully excluded the trials in which the processing task started with a free delay, and the processing patterns in which there was not a constant number of distractors presented consecutively, as these may possibly reduce memory performance for other reasons than switching costs.

In the same vein, the comparison exploring the irregularity factor (i.e. 11001100 *vs.* 11100010) examined patterns with no consolidation opportunity and no switch variations. Recall, that the regular patterns have a constant number (two) of distractors presented consecutively and a fixed duration of free delays of 1200. In contrast, the irregular pattern has a variable amount of distractors in a row (three or one) and a variable delay of free time (1800 ms or 600 ms).

Finally, the consolidation hypothesis was explored by comparing patterns with and without a window of free time at the beginning of the processing task, while maintaining constant the number of switches and the (ir)regularity of the patterns. In other words, we conducted the following pairwise comparisons: 01010101 *vs.* 10101010 and 00110011 *vs.* 11001100 and 00011101 *vs.* 11100010, with respectively a consolidation opportunity of 0 ms *vs.* 600 ms or 1200 ms or 1800 ms.

For all pairwise comparisons, we conducted separate analysis using the same five statistical models as previously (see Table 3.5 for a summary of the models).The results showed for all comparisons similar findings to our main analysis (list length model $M1$ ranked best and the processing pattern model $M2$ ranked last), except for 00110011 *vs.* 11001100 (with respectively a window of free time of 1200 ms *vs.* 0 ms before the presentation of the first distractor).

Table 3.6 shows the ranking of the models when the analysis included only the consolidation processing patterns (i.e. 00110011 *vs.* 11001100). The top-rank model was the full model without interaction ($M3$). When breaking down the results, we saw strong evidence in favor of the contribution of list length to memory performance ($M1/M0 > 150$), whereas the consolidation model hardly distinguished from the null model ($M2/M0 = 2.47$). The lack of clear influence of a delayed processing pattern on memory performance could explain why the AICcs of the first three statistical models are so close to one another: evidence suggested support for the full models, but we could not discard the simpler model incorporating only list length ($M3/M1 = 2.72$; $M3/M4 = 1.14$).

Beyond the effect of memory load already reported, the best ranked model showed no effect of the delayed pattern on memory performance when the Bonferroni correction was applied ($F(1, 645.12) = 4.04$, $p = .045$). The estimate showed an increase of memory performance of only 2.5% when a free delay of 1200 ms was provided at the beginning of the processing task.

In summary, our results did not provide evidence in favor of an influence of the variations of the processing schedules on memory performance, whether the patterns included switching, irregularity or consolidation. However, examination of memory performance showed that 87% of the trials were perfectly recalled by the participants. Table 3.7 shows the mean memory performance as a function of the processing patterns. As in our previous experiment, the recall task was still likely too easy. Therefore, it is possible that the effects of the processing schedules was again not apparent because of this ceiling effect. We attempted to let the models better predict the partially incorrect responses by excluding the trials that were either perfectly recalled or completely failed, but nonetheless, the pattern of statistical results shown in Table 3.6 remained.

Table 3.7: Mean memory performance (SD) as a function the processing patterns.

00011101	00110011	01010101	11001100	11100010	10101010
.94 (.16)	.96 (.12)	.95 (.13)	.93 (.15)	.95 (.12)	.94 (.14)

Analysis of concurrent task performance. We used the proportion of correctly executed decisions in the concurrent task within a trial as our dependent variable (i.e. sensitivity). We proceeded to model comparison following the same rationale as in our previous analysis (see Table 3.5 for a summary of the statistical models). Table 3.8 shows a ranking of the models. The statistical analysis indicated that the null model was ranked as the most likely model given the observations, indicating that the effect of processing patterns on the concurrent task performance was null. As in our previous analysis, we pursued our analysis to examine specifically the switch factor (i.e.11001100 *vs.* 10101010), the irregularity factor (i.e. 11001100 *vs.* 11100010) and the consolidation factor (i.e. 01010101 *vs.* 10101010 and 00110011 *vs.* 11001100 and 00011101 *vs.* 11100010). The results showed for all comparisons similar findings to our main analysis, that is, the null model was ranked best.

Table 3.8: Summary of model selection for the concurrent task in Experiment 3 based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 3.5.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M0$	3	-874.99	0.00	0.39
$M2$	8	-874.41	0.58	0.29
$M1$	4	-873.43	1.56	0.18
$M3$	9	-872.84	2.16	0.13
$M4$	14	-864.82	10.17	0.00

3.2.3 Discussion

The goal of the present study was to investigate the effect of processing schedule variations on working memory. We assumed that these schedules could cause either beneficial or detrimental effects depending on how the processing events are displayed within a trial. Specifically, we delineated three predictions: switching and irregularity costs should impair recall and concurrent task performance, whereas consolidation should have the opposite effect and enhance performance.

Overall, our results do not confirm our hypothesis. However, we observed a ceiling effect on recall performance, which might explain the lack of effectiveness of our manipulation. Although we reduced the list lengths of the trials in this study compared to Experiment 2, the ceiling effect is probably not a result of insufficient memory load alone. In all likelihood, the cognitive load of our task is also at fault. Our concurrent task is similar to the continuous performance task paradigm, which is commonly known to measure selective and sustained attention/vigilance (Roebuck, Freigang, & Barry, 2016). There are many variations of the task used in the literature, but often the display duration of the stimuli are shorter (varying from 40 to 200 ms), while the interstimulus intervals are longer, ranging typically from one to two seconds (van den Bosch, Rombouts, & van Asma, 1996). Since our goal was not to measure attention per se, but simply to distract it away from storage activities, we adapted this paradigm to a complex span task. We believed that the typically used display duration would make the task too difficult for the participants, but the behavioral data suggested otherwise. Therefore, it is possible that in our experiment the relatively slow pace of the presentation of the distractors in regards of its attentional demand has reduced the cognitive load of the task to a point where all participants were able to recall almost every letter.

To summarize, the results of Experiment 1 and 2 do not support the idea that the structure of a task can produce a detrimental effect of switching on complex span task performance. However, the methodological issues aforementioned cast doubts on our experimental design. In the following section, we will pursue in this direction with a task associated to a higher cognitive load in the hope to create the preconditions for the effectiveness of our manipulation.

3.3 Experiment 4

This study investigated the same hypothesis as in Experiment 3, but with a task that had a purported higher cognitive load. Effectively, our previous experiment showed a ceiling effect in performance, which might explain why our manipulation of processing patterns was not effective. To maximize our chances to observe switching costs in working memory performance while still favoring the use of elementary processing steps, we opted for a reading digit span task that has been extensively used by the authors of the TBRS model. We inserted the same processing schedules as in Experiment 3 in our complex span task, and expected again an effect of switching, consolidation and irregularity of the processing patterns on performance.

3.3.1 Method

Participants. Eighty-three students of Université Côte d'Azur participated to this experiment in exchange for partial course credit (70 females, 13 males; mean age = 22.5 years, $SD = 5.0$). All participants were tested in a quiet room at the university. The duration of the experiment was approximately one hour.

Material and procedure. Participants were invited to memorize letters while doing a concurrent task. The storage and processing items (i.e. letters and digits) were identical to Experiment 2 and 3. However, this time the participants were instructed to read aloud every digit and letter displayed. Additionally, in order to collect reaction times, the participants had to press on the space bar at the same time they read an item aloud. The patterns of the secondary task were identical to Experiment 3 (Figure 3.5). The length of the trials was progressive from 1 to 6 letters, with 12 repetition per set size. For each set size every pattern was repeated twice. The order of the patterns was randomly drawn without replacement.

Each trial started with a fixation cross displayed for 750 ms, after which the letter appeared for 1500 ms. After the consonant, the processing events (i.e. four digit and four slot of free time) were displayed, each for 600 ms. A second computer screen was plugged to the computer of the participant enabling the monitoring of the oral responses (i.e. reading aloud). This second screen was placed on another table behind the participant at a distance of approximately 2 meters. At the end of each trial, the participants were instructed to recall the letters in the correct serial order. The recall procedure was identical to Experiment 3. Practice trials of about 10 minutes were performed by the participants before the beginning of the experiment. This training consisted of six trials of one letter and six trials of two letters each with different patterns of the processing task.

3.3.2 Results

Results from two participants were excluded from analysis as the proportion of correctly recalled items deviated more than two standard deviation from other participants. For the remaining 81 participants (mean age = 22.2 years, $SD = 4.4$). We used the same scoring method as in our previous experiments (i.e. proportion of correctly recalled items in their correct serial position), referred as memory performance from now on.

Analysis of the recall task We re-conducted the same analysis as in Experiment 3 (see Table 3.5 for a summary of the statistical models).

Table 3.9 shows a ranking of the models in terms of of the best fit, and the details of the number of parameters of each model (K), the AICc values, the delta AICc values and the weight of each model. The results showed that the enhanced model incorporating list length and the processing patterns with the interaction term ($M4$) was best ranked, and considered as the most likely model given the observations. As seen the top ranking model has an Akaike weight of 1, and consequently the calculated evidence ratios suggest all very strong evidence in favor of the full model with interaction, we can thus reject the alternative models ($M4/M3 > 150$; $M3/M1 > 150$).

Regarding the parameters of interest, we observed an effect of memory load, with higher performance on the recall task for shorter list lengths compared to longer list lengths ($F(1, 5680.0) = 2158.4$, $p < .001$). The estimate showed a -7% memory performance decrease when memory load was increased by 1 unit. The main effect of processing patterns on memory performance showed no significant difference ($F(5, 5680.3) = 0.9396$, $p = 0.45$). However, the interaction between memory load and the processing patterns produced significant differences on memory performance ($F(5, 5680.4) = 10.3973$, $p < .001$). Figure 3.7 shows the mean memory performance as a function of the rhythmic patterns and memory load. As seen from the figure, the manipulation of processing schedules is not apparent for smaller set sizes, but we could observe a decrease on memory performance as a function of the processing pattern as memory load increased ($L > 4$). In sum, the result showed that our manipulation of processing schedules is effective (provided that storage demand is high). In the following, we proceeded to pairwise comparisons as in Experiment 3 to assess to what extent switching, consolidation and irregularity of the processing patterns contributed to differences in memorization.

Effect of switching on memory performance. We now report the results of the processing patterns which allowed manipulation of the switch factor (i.e.11001100 *vs.* 10101010, requiring three and seven switches respectively).

Table 3.9 shows the ranking of the models. The results showed that when the analysis included only the switch patterns, the best model is the one incorporating list length and the switch factor without the interaction term ($M3$). Breaking down the results, we saw strong evidence in favor of the contribution of memory load to memory performance ($M1/M0 > 150$). This is not the case for the switch factor ($M2/M0 = 1.11$), meaning that the influence of switching on memory performance is minimal at best. This explains why the statistical ranking shows equal support for the first three models ($M3/M1$ = 2.08; $M3/M4$ = 2.15).

Regarding the parameters of interest of the model best ranked, we observed an effect of memory load, with higher performance on the recall task for shorter list lengths compared to longer list lengths ($F(1, 1831.4) = 878.3$, $p < .001$). The estimate showed a -10% memory performance decrease when memory load was increased by 1 unit. The main effect of switching on memory performance showed no significant differences ($F(1, 1831.3) = 3.5$, $p = 0.06$). The estimate showed a -1.5% memory performance decrease when the number of switches was increased by 1 unit.

Table 3.9: Summary of model selection in Experiment 4 for memory performance based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 3.5.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Data including all processing patterns				
$M4$	14	1471.91	0.0	1
$M3$	9	1513.61	41.71	0.0
$M1$	4	1616.19	144.28	0.0
$M2$	8	3326.14	1854.23	0.0
$M0$	3	3398.30	1926.40	0.00
Switch factor, no consolidation, no irregularity (11001100 *vs.* 10101010)				
$M3$	5	649.7	0.0	0.51
$M1$	4	651.16	1.47	0.25
$M4$	6	651.23	1.53	0.24
$M2$	4	1365.14	715.44	0.0
$M0$	3	1365.34	715.65	0.00
Consolidation factor, no irregularity, no variation of switches (00110011 *vs.* 11001100)				
$M4$	6	614.81	0.0	0.98
$M3$	5	623.01	8.20	0.02
$M1$	4	638.30	23.49	0.0
$M2$	4	1197.2	583.12	0.0
$M0$	3	1208.45	593.64	0.00
Irregularity factor, no consolidation, no variation of switches (11001100 *vs.* 11100010)				
$M1$	4	687.68	0.0	0.57
$M3$	5	688.89	1.20	0.31
$M4$	6	690.88	3.20	0.12
$M0$	3	1374.57	686.89	0.0
$M2$	4	1376.19	688.51	0.00

*Note: M0 = Null model; M1 = List Length ;M2 = Processing pattern; M3 = List length + Processing pattern; M4 = List Length * Processing pattern.*

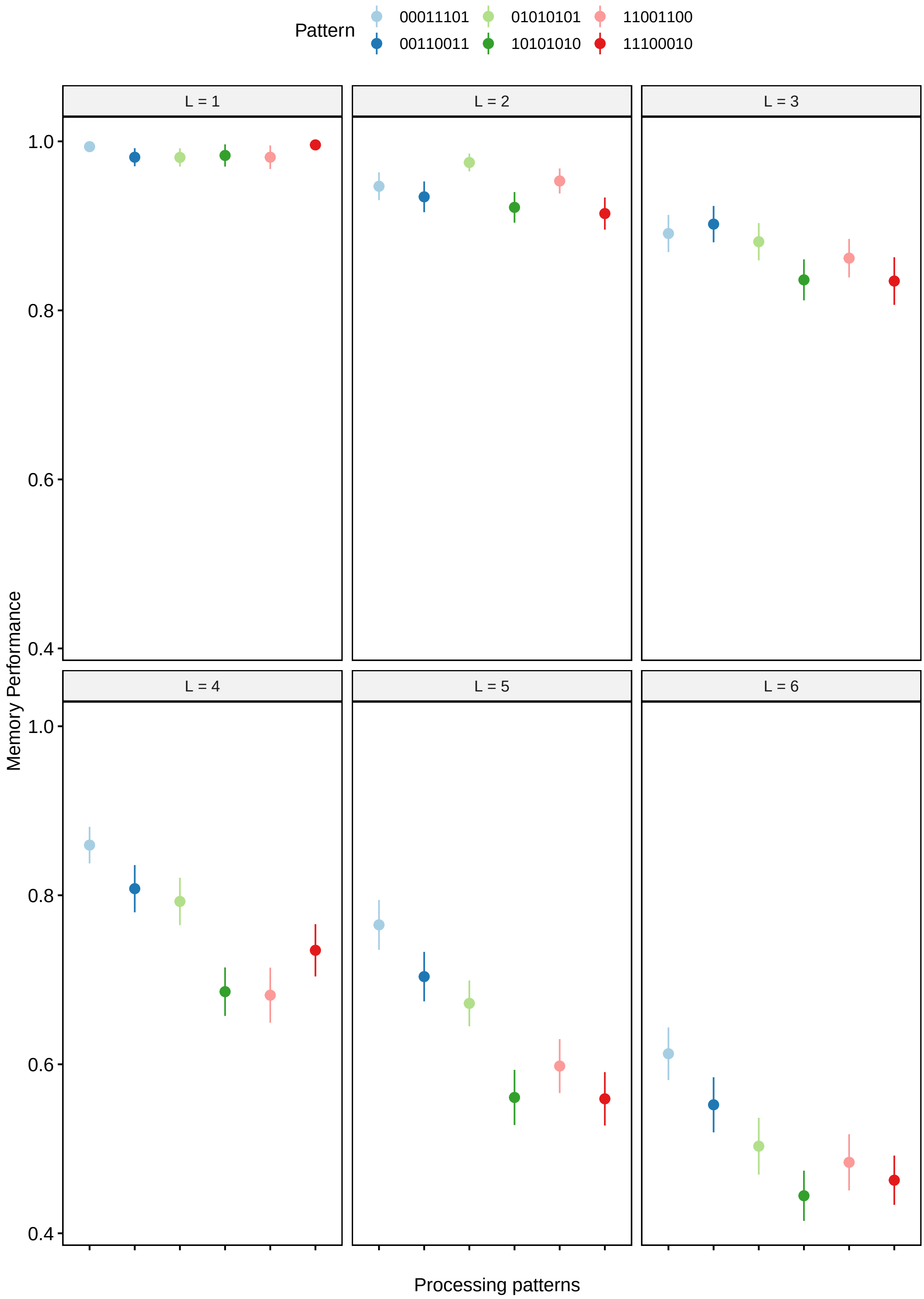


Figure 3.7: Mean memory performance as a function of processing patterns of the concurrent task in Experiment 4, split by list length (from 1 to 6 letters).

Effect of consolidation on memory performance. In this section, we report the results of the processing patterns manipulating only the consolidation factor (i.e. 00110011 *vs.* 11001100, 01010101 *vs.* 10101010, and 00011101 *vs.* 11100010).

Table 3.9 shows a ranking of the models when the analysis included only consolidation patterns. In essence the analysis yielded similar results than when all the patterns were examined together ($M4$ best ranked), but the full model without the interaction term is not to be discarded yet ($M4/M3 = 60.47$). These results were replicated when the pairwise comparison involved the processing pattern with a shorter consolidation opportunity (600 ms) and consequently a higher number of switches in both patterns (i.e. 01010101 *vs.* 10101010) and, also when the processing pattern involved a longer consolidation opportunity (1800 ms) and irregular processing events (i.e. 00011101 *vs.* 11100010).

Regarding the parameters of interest of the model best ranked, we observed an effect of memory load, with higher performance on the recall task for shorter list lengths compared to longer list lengths ($F(1, 1840.6) = 266.2$, $p < .001$). The estimate showed a -8% memory performance decrease when memory load was increased by 1. The main effect of consolidation processing patterns on memory performance showed no significant differences ($F(1, 1841.1) = 1.1$, $p = 0.30$). The estimate showed a 2% increase of memory performance as an opportunity for consolidation was provided. The interaction between list length and consolidation patterns showed a significant effect on memory performance ($F(1, 1841.1) = 10.25$, $p < .001$).

Effect of irregularity of processing patterns on memory performance. In the following, we report the results of the processing patterns manipulating only the irregularity factor (i.e. 11001100 *vs.* 11100010).

Table 3.9 shows a ranking of the models. The results showed that $M1$ was the best ranked model. Evidence in favor of the model incorporating only list length compared to the null model is strong ($M1/M0 > 150$). Moreover, the irregularity processing pattern model ($M2$) was ranked last after the null model which indicates that the explanatory power of this factor is nonexistent. Regarding the parameters of interest, we observed unsurprisingly a significant effect of memory load on memory performance ($F(1, 1842.4) = 835.25$, $p < .001$), with an estimate showing a -10% memory performance decrease when memory load is increased by 1 letter.

Analysis of RTs. We conducted the same analysis as in the previous section to examine the effect of processing patterns on reaction times of the reading digit task. Our dependent variable was the mean reading reaction times (RTs) in ms trial-by-trial (invalid RTs were excluded from analysis).

Table 3.10 shows a ranking of the models when the dependent variable was the mean RTs. The results showed that the full model with the interaction term ($M4$) was best ranked. The calculated evidence ratio of the two models including only one factor of interest was found very strong ($M1/M0 > 150$; $M2/M0 > 150$), which indicates that both memory load and the number of switches contributed together to the RTs of the task. The calculated evidence ratios between the full models with and without the interaction suggested also strong evidence for the model including the interaction term ($M4/M3 > 150$).

Regarding the parameter of interest of the model best ranked, both main effect showed a significant effects on RTs ($F(1, 5628.0) = 131.02$, $p < .001$; $F(5, 5628.1) = 62.48$, $p < .001$ for list length and the processing pattern factor respectively). The interaction between list length and processing patterns was also significant ($F(5, 5628.1) = 5.0$, $p < .001$). The estimate showed a decrease of RTs of 3 ms when memory load is increased by 1. Regarding the processing pattern estimates, in reference to pattern 00011101 showed a decrease of RTs of 6 ms, 47 ms , 25 ms when the task involved the patterns 00110011, 01010101, 10101010 respectively and an increase of 2 ms and 4 ms for the pattern 11001100 and 11100010. Figure 3.8 shows the mean RTs of the reading digit task as a function of list length and the processing schedules of the task. As seen, the fastest RTs are associated to patterns with seven switches, and this intensifies as the length list increases. As in the previous section, we proceeded in the following to pairwise comparisons by examining specifically the switching, consolidation and irregularity factor.

Effect of switching on RTs. The results of the next section concerns the pair-wise comparison of the processing patterns manipulating only the switch factor (i.e.11001100 *vs.* 10101010, with respectively three and seven switches). Ranking of models is showed in Table 3.10. The best ranked model included list length and the switch factor without the interaction term ($M3$). We observed strong evidence in favor of both factors of interest ($M2/M0 > 150$; $M1/M0 > 150$), meaning that both memory load and switching contributed to the full model. Evidence suggested equal support between the enhanced models (with and without the interaction term ($M3/M4 = 1.79$).

Table 3.10: Summary of model selection in Experiment 4 for RT data of the concurrent task based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 3.5.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Data including all processing patterns				
$M4$	14	56737.56	0.00	1
$M3$	9	56752.41	14.85	0.00
$M2$	8	56879.62	142.07	0.00
$M1$	4	57897.72	1160.17	0.00
$M0$	3	58003.37	1265.82	0.00
Switch factor, no consolidation, no irregularity (11001100 *vs.* 10101010)				
$M3$	5	18879.63	0.00	0.64
$M4$	6	18880.79	1.17	0.36
$M2$	4	18966.57	86.95	0.00
$M1$	4	19208.76	329.13	0.00
$M0$	3	19281.85	402.23	0.00
Consolidation factor, no irregularity, no variation of switches (00110011 *vs.* 11001100)				
$M4$	6	18944.80	0.00	0.44
$M3$	5	18944.82	0.01	0.44
$M1$	4	18947.31	2.50	0.13
$M2$	4	18995.35	50.55	0.00
$M0$	3	18997.77	52.97	0.00
Irregularity factor, no consolidation, no variation of switches (11001100 *vs.* 11100010)				
$M3$	5	19190.65	0.00	0.56
$M4$	6	19192.41	1.76	0.23
$M1$	4	19192.59	1.94	0.21
$M2$	4	19245.51	54.86	0.00
$M0$	3	19247.41	56.76	0.00

*Note: M0 = Null model; M1 = List Length ;M2 = Processing pattern; M3 = List length + Processing pattern; M4 = List Length * Processing pattern.*

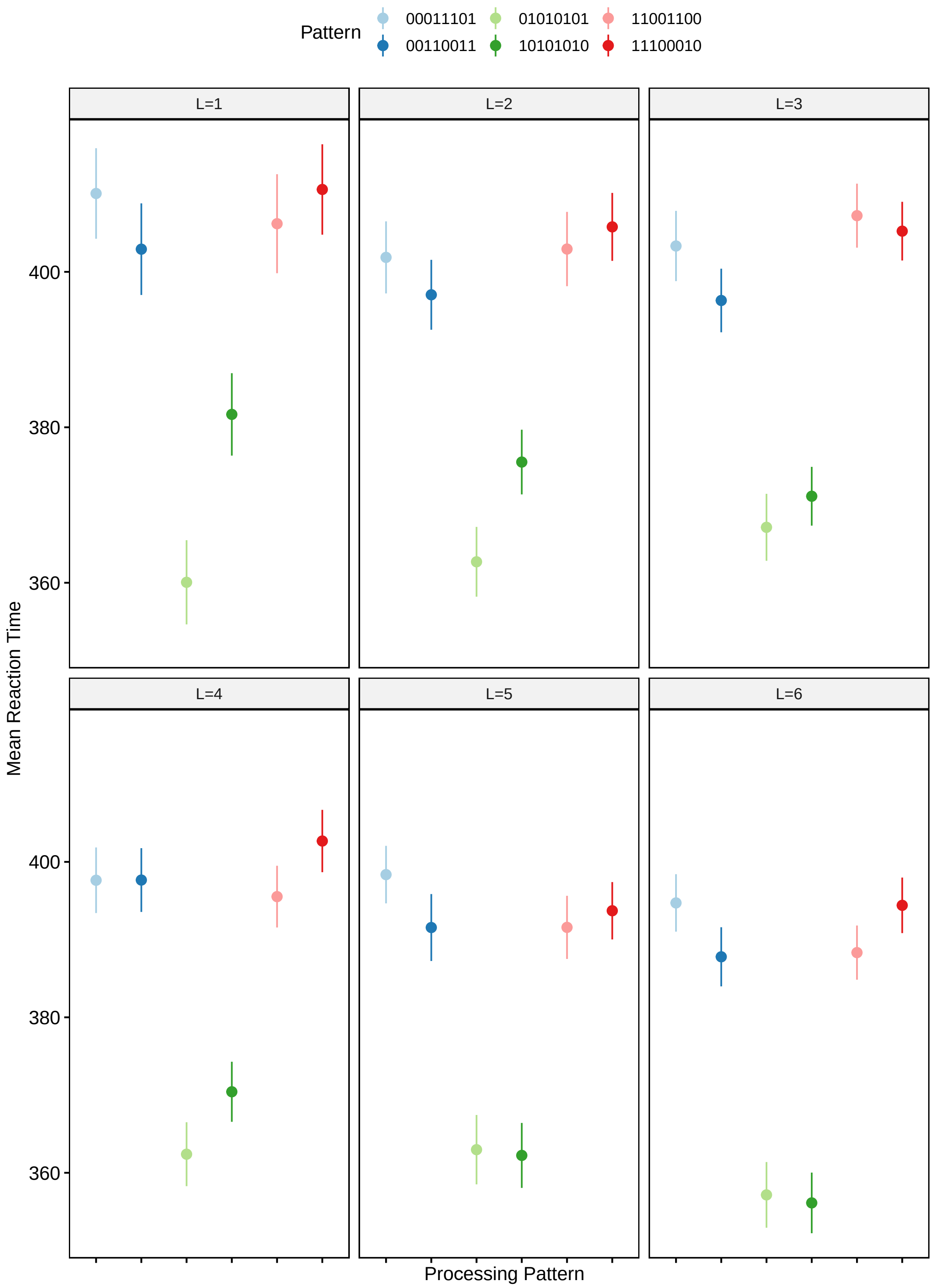


Figure 3.8: Mean RTs (ms) of the reading digit task as a function of the processing patterns in Experiment 4, split by list length (from 1 to 6 letters).

Regarding the model best ranked, we observed an effect of memory load and the number of switches on RTs ($F(1, 1810.2) = 91.18$, $p < .001$; $F(1, 1810.4) = 363.43$, $p < .001$). The estimate showed a decrease of 4 ms when memory load was increased by 1. The switching estimate showed a decrease of 30 ms when the task went from three to seven switches.

Effect of consolidation on RTs. The results in the next section concerns the consolidation processing patterns (i.e. 00110011 *vs.* 11001100, with a consolidation window respectively equal to 1200 ms *vs.* 0 ms). Ranking of models is showed in Table 3.10.

The best model is the full model with the interaction term. A closer look at the result showed that both factors of interest contributed separately to RTs, although evidence was less convincing for the consolidation factor ($M1/M0 > 150$; $M2/M0 = 3.36$). We then compared the full model without the interaction term and the model including only the list length factor. The comparison favored the former model, but did not allow to discard the latter one ($M3/M1 = 3.48$). Finally, we observed equal support for both enhanced models with or without the interaction term ($M4/M3 = 1.01$).

Regarding the parameters of the model best ranked, we observed an effect of memory load ($F(1, 1819.1) = 53.71$, $p < .001$), with an estimate showing a decrease of 3 ms when memory load was increased by 1. The main effect of consolidation opportunity and its interaction with memory load on RTs were not significant when Bonferroni corrections were applied (all ps>0.16).

When the comparison involved the 0101010 *vs.* 10101010 patterns (i.e. 600 ms *vs.* 0 ms consolidation window), similar results were found with stronger support for the top-ranked model ($M4$). Hence, both the main effect of the consolidation pattern and its interaction with the processing pattern were statistically significant ($F(1, 1820.2) = 43.026$, $p < .001$; $F(1, 1820.3) = 24.366$, $p < .001$ for the main effect and the interaction respectively). In this case, the consolidation estimate showed an increased of 22 ms of the RTs when no consolidation was provided. As already seen in Figure 3.8, as the list length increases, the reaction times of the pattern without a consolidation opportunity decreases. In contrast, comparison of the patterns 00011101*vs.* 11100010 (i.e. delay of free time of 1800 ms *vs.* 0 ms) indicated no effect of consolidation on RTs (top-ranked model $M1$ and $M2$ ranked last).

Effect of irregularity on RTs. The results in the next section concerns the regularity processing patterns (i.e. 11001100 *vs.* 1110001). Again, ranking of models is showed in Table 3.10.

The best ranked model was the full model without the interaction term ($M3$). Evidence suggested equal support for the first three models ($M3/M1 = 2.64$; $M3/M4 = 2.41$). We observed strong evidence in favor of the influence of memory load on RTs ($M1/M0 > 150$). In contrast, the influence of irregularity on RTs appeared minimal ($M2/M0 = 2.64$).

Regarding the parameters of interest, we observed an effect of memory load ($F(1, 1818.3) = 57.77$, $p < .001$). The estimate showed a decrease of 4 ms when memory load was increased by 1. The effect of irregularity was not significant when Bonferroni correction was added. The irregularity factor estimate showed an increase of 3 ms when the task was executed with an irregular processing schedule.

3.3.3 Computational simulations of processing patterns

The TBRS model assumes that working memory performance is determined by the cognitive load of the task at hand. Therefore, the model does not predict any differences on recall performance as a function of the processing schedules used in this experiment. In order to fathom with more precision the predictions of the verbal model, computational simulations of the TBRS* and the TBRS2 model were carried out. The simulation of the TBRS* was implemented by Benoit Lemaire and Sophie Portrat (personal communication, 2018). As to the simulation of the TBRS2 model, it was conducted by Nicolas Gauvrit.

Regarding the TBRS* simulations, all presentation durations of storage and processing events were provided as input to the TBRS* model (version 3.2.4). The reaction times for processing a distractor allowed to calculate the duration of the attentional capture for each distractor, and by extension the average cognitive load for each pattern was computed. This was achieved by dividing the RT's by the duration of the distractors. Table 3.11 shows the mean memory performance and cognitive load as a function of the processing patterns of participants and the prediction of both the TBRS* and TBRS2 model.

Table 3.11: Summary of mean memory performance of participants and modeling predictions and mean cognitive load (CL) as a function the processing patterns.

	00011101	00110011	01010101	11001100	11100010	10101010
Participants	.872	.837	.819	.810	.804	.756
TBRS*	.808	.804	.858	.847	.811	.873
TBRS2	.886	.888	.815	.874	.876	.788
CL	.315	.312	.275	.323	.326	.298

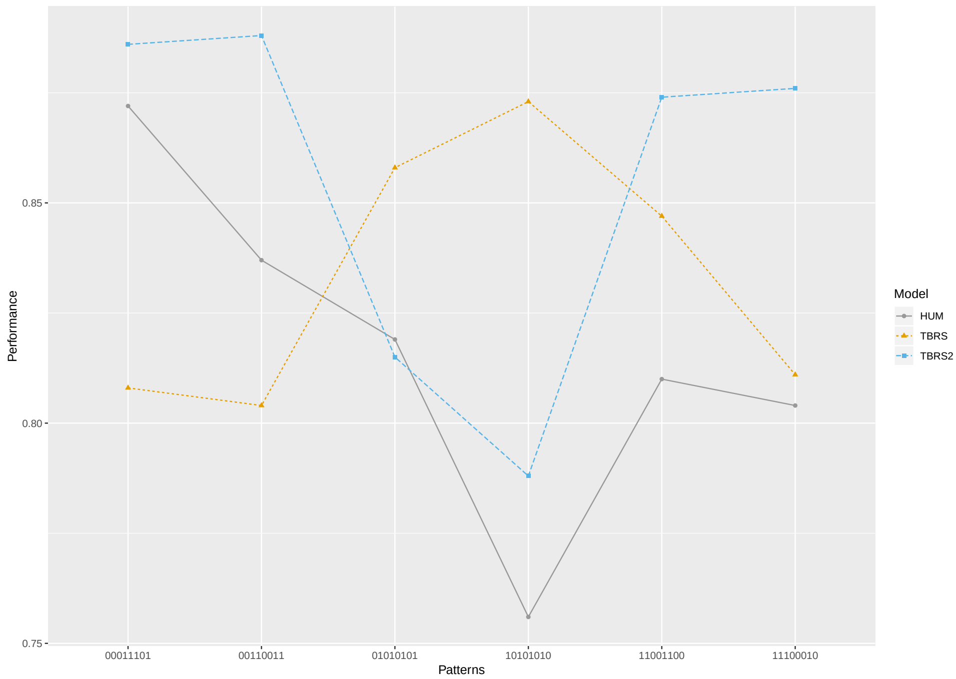


Figure 3.9: Mean memory performance of participants and model simulations as a function of the processing patterns of Experiment 4. *Note*: The TBRS* simulation is based on fixed parameters.The TBRS2 parameters are the following: baseline = 3, sw = 20.

Figure 3.9 shows the estimation of the two computational models compared to the experimental data. As seen, the simulation of the TBRS* model does not account very well for the data observed. Best memory performance was expected to be associated to the processing patterns 10101010, but the exact opposite was observed in humans. On visual inspection the TBRS2 seems to account better for the human data than the TBRS* model. Regarding the simulation of the TBRS2, the simulation was conducted with an optional parameter sw, which takes into consideration the time to switch from storage to processing (and the other way around). In other words, the model implements switch costs by simply capturing attention for an additional duration even though the distractor has already been processed. Additionally, the TBRS2 does not need a specific parameter to account for the consolidation process. Despite the absence of a parameter for consolidation, the TBRS2 is able to explain small variations in probability of recall. This phenomenon will be discussed in Chapter 5.

3.3.4 Discussion

The present study aimed to examine the effect of variations of processing schedules on working memory performance. Our results indicates that the processing pattern may influence working memory performance. For the recall task, our findings suggest evidence in favor of consolidation in working memory. In contrast, switching and irregularity of the processing patterns did not affect memory performance. For the concurrent task, our results indicated a beneficial effect of switching on reaction times of the reading digit task, whereas evidence in favor of consolidation was more elusive. The irregularity of the patterns did not impact concurrent task performance.

1) Consolidation & working memory performance

According to the consolidation hypothesis, better memory performance is associated with the presentation of a free delay right after the presentation of a to-be-remembered stimulus. In contrast, when a distractor is immediately presented after the stimulus, reduced memory performance is expected. Regarding the concurrent task, the consolidation hypothesis predicts slower response times associated with the interruption of the consolidation process. This is because the presentation of a distractor immediately after the stimulus is thought to interrupt consolidation processes, which leads to slower reaction times.

Our results show that memory performance is positively influenced by a delay of free time before the presentation of the first distractor. This benefit is, however, only observed when the storage demand is high. The requisite of high memory load for consolidation effects to be unveiled seems intuitive. For short list lengths, absence of consolidation could simply be due to a ceiling effect, whereas for longer list lengths the effect is probably caused by the inherent capacity limitations of working memory. Nonetheless, our findings suggest that consolidation appears to have only a small impact on memory when the list length effect is taken into account, with an increase of only 2% of memory performance when a delay of 1200 ms is provided before the presentation of the first distractor. In line with our findings, Bayliss et al. (2015) and later De Schrijver and Barrouillet (2017) had established the presence of consolidation effects in complex span tasks (although the magnitude of the effect in their study was more important than in ours). A possible explanation for this slight discrepancy is that the display duration of the storage items in this experiment was much longer than typically used in working memory consolidation studies. Each to-be-recalled letter was presented for 1500 ms, while it is generally at most 500 ms in other studies. Therefore, it is possible that post-encoding processes (i.e. consolidation) were well underway by the time the event (a delay or a distractor) occurred in our study.

In line with our finding that consolidation is memory load dependent, some studies have failed to find a beneficial effect of consolidation opportunity on recall performance when using low demanding tasks. For instance, in a dual task paradigm, Jolicœur and Dell'Acqua (1998) observed slower completion of the secondary task associated to shorter consolidation times, while recall performance remained unaffected. In their experiments participants had to remember at most three items. In contrast, in a similar dual task paradigm using a higher memory load and a more demanding secondary task, Nieuwenstein and Wyble (2014) were able to establish such a link between memory performance and consolidation, with better performance associated to longer opportunities of consolidation. This was interpreted as more direct evidence in favor of consolidation in short-term memory.

Although, longer delays of free time immediately after the presentation of the storage event are associated in some cases to better memory performance, we argue that at least two alternative explanations might account for the same phenomenon without the need to imply any additional processes in working memory. Firstly, it could be that any maintenance ventures are simply more beneficial to memory performance when they occur before a processing interruption rather than after.

This alternative has also been discussed by Bayliss et al. (2015), although it was discarded by the authors. Secondly, consolidation effects could be a by-product of a cumulative refreshment schedule, and such effects could be expected when all memory traces are not equally refreshed. In line with this reasoning, the authors of the TBRS model (Barrouillet & Camos, 2014a), specify that one subtle variation of the TBRS is that it is not always possible to sufficiently refresh all of the memory items. We will address this question in more depth in Chapter 5.

Regarding the secondary task, the RT data provides ambiguous results. In agreement with the consolidation hypothesis, our results showed faster reaction times associated with the presentation of a free delay after the storage episode. However, our finding is limited only to the shortest consolidation time (600 ms), although there was still a trend for faster responses when consolidation duration was 1200 ms. As discussed above, one reason for this might be that post-encoding processes are already completed before the presentation of the first distractor. However, this seems an unlikely explanation, since response times are faster as the memory load increases, and particularly so when an immediate distractor is presented. If anything, we should have observed the opposite trend according to the consolidation hypothesis. Logically, consolidation of several items should be more time consuming than the consolidation of a few items. Thus, the immediate presentation of a distractor should be especially detrimental in high memory load conditions, leading thus to slower processing [2]. The behavioral data suggest, however, otherwise.

One possible explanation to account for our RT results is related to the idea that preparation time (i.e. time taken before switching from storage activities to processing a distractor) can lead to faster execution of the task. Preparation to the upcoming task is known to reduce response times to execute the task, and even more so in task-switching trials compared to task-repetition trials (e.g. Koch, 2003; Meiran, 1996; Monsell & Mizon, 2006). However, task preparation does not allow to explain why response times are faster as list length (or memory load) increases or why the pattern that has the shortest delay of free time after the presentation of the storage item has the fastest response times.

Accounting for this phenomenon requires to consider that other processes are involved perhaps along with task preparation or consolidation. Closely related to

[2]This argument holds only if we assume that consolidation affects all the storage items that have entered working memory so far in a given trial. However, it could be argued that consolidation processes affect only the last storage item that has been encoded. In this case, RTs should not vary as a function of the memory load.

the notion of preparation of the upcoming task is the predictability of the task. Task predictability is thought to support the preparation of the upcoming task (Nicholson, Karayanidis, Davies, & Michie, 2006). According to Vandierendonck et al. (2007), predictable sequences give additional opportunities for preparation and learning. Since in our experiment, the processing pattern used in a given trial was repeated between the letters to remember (i.e. for three letters the participant completed the task with the same pattern three times: L10101010L10101010L10101010), it is fair to assume that the processing schedule is more predictable in longer list length as the participant can progressively detect and learn the pattern throughout the subsequent repetitions. Moreover, the short consolidation pattern and its symmetrical version (i.e. 01010101 *vs.* 10101010) are the two most predictable processing schedules. Once the pattern is learnt participant can predict the next events as soon as the first two elements are presented, whereas this requires the presentation of three subsequent event for all the other processing schedules. Taken together, preparation, predictability and consolidation do not allow to explain in a simple manner these results. Thus, their exact meaning remains unclear for the moment.

2) Switching & working memory performance

According to our switching hypothesis, when distractors and blank delays are presented in a sequence one after the other, the numerous alternations between storage and processing events should induce switching costs potentially impairing memory performance. Against our hypothesis, and in line with previous work (e.g. Logan, 2004), our results show that the number of switches prompted by the structure of the task does not influence recall performance. We could observe a slight trend in favor of switching costs on memory performance, but this trend was not statistically significant.

Regarding the effect of switching on the concurrent task, our results show a beneficial effect of the number of switches on reaction times. This is at odds with our hypothesis, since we assumed that switching would cause the opposite effect. Neither consolidation nor the task-preparation hypothesis can easily explain these results, since the considered switching patterns involved the presentation of a distractor immediately after the storage event. As discussed above, we can emit the hypothesis that task predictability could have benefited the patterns with a high number of switches (i.e. 10101010), because they can be easily learnt and recognized. In this case, switch costs could perhaps be neutralized.

Overall, neither memory performance nor performance at the concurrent task favor our hypothesis. We believe that this question still merits attention, and increasing the difficulty of the concurrent task could perhaps bring to light a clearer pattern of results, with the idea that a difficult task could increase switching costs. This question will be examined further in Chapter 4.

3) Irregularity & working memory performance

Our last hypothesis explored the potential effect of irregular processing patterns. The idea was that patterns that do not present a constant number of distractors and have variable interstimulus intervals between the distractor should reduce recall performance.

Neither memory performance nor concurrent task performance showed evidence of an effect of irregularity of the processing patterns. One possible reason for this result might be that once the participants learned the different processing patterns by repetitions throughout the list length, as discussed above, the awareness of the irregularities in the sequence might have reduced potential irregularity costs. Experiment 5 will explore further this idea.

3.4 Experiment 5

We now make a distinction between the regularity of a pattern in the concurrent task and the predictability of the patterns. The patterns can be regular (for instance 01010101 but not predictable, if the participant cannot anticipate they will get this pattern in particular). On the contrary, a pattern can be considered less regular (for instance 01000111) but participants could know they will get this pattern. The aim of the following experiment was to examine the effect of an unpredictable task structure on working memory performance. In our previous experiments, we examined whether we could identify a switching cost in complex span task performance. This research question was studied by manipulating the number of switches to perform between storage and processing events. However, because we used the same processing patterns between each letter to remember, it is possible that the participants were able to prepare for a task switch once they felt familiar with the pattern in question. Such an anticipation by the participants might have reduced potential switch costs.

For this experiment, we made the hypothesis that an unpredictable task structure should have a detrimental effect on working memory performance. Participants should not be able to prepare locally for a task switch when the distractors in the concurrent task are unpredictable. In contrast, *if* the concurrent task became predictable globally as in Experiment 3 and 4 (i.e. again, as the same processing schedule repeated throughout the trial), participants could have performed better locally (i.e. as they might have been able to anticipate refreshment opportunities and prepare for a task switch accordingly).

In order to assess this issue of global task predictability on working memory performance, we used a complex span task with the same types of processing patterns as in our previous experiment. The main difference was that this new experiment made use of different patterns between to-be-recalled items, within trials, when the task was not predictable.

3.4.1 Method

Participants. Forty-four students of Université Côte d'Azur participated to this experiment in exchange for partial course credit (35 females, 9 males; mean age = 22.0 years, $SD = 4.3$). All participants were tested in a quiet room at the university. The duration of the experiment was approximately one hour.

Material and procedure. We used in this study a reading digit span task, with identical storage and processing stimuli as in Experiment 4 (i.e. letters and digits). All participants underwent two experimental conditions presented in a block design.

In the predictable condition, participants executed all trials built with a unique processing pattern. For half of the participants the unique pattern started with a distractor: the chosen pattern was 11001100. For the other half, a delay of 1200 ms was presented before the presentation of the first distractor: the pattern was 00110011. Recall that 1 stands for the presence of a to-be-processed digit and 0 for a free delay. The 1200 ms delay is represented by the two zeros at the beginning of the pattern 00110011. In sum, this condition was predictable because a unique pattern was used repetitively and because the pattern alternated regularly. This manipulation maximized the chance for the participant to capture the regularity and use it to their benefit.

In the unpredictable condition, the processing patterns of the the concurrent task always changed between the letters to remember. For instance, in a sequence of three letters to be remembered, three different processing patterns were presented. The processing patterns were drawn randomly without replacement from a pool of 10 different patterns, each being presented a total of 14 times throughout all trials. Half of the processing patterns used in this condition started immediately with a distractor after the presentation of a to-be-remembered letter, and the other half with a delay of free time. Depending on the chosen pattern, the delay of free time varied from from 600 ms to 1800 ms (before the first digit occurred after a letter). Simply, this refers to the presence of one, two, or three zeros at the beginning of our patterns. Figure 3.10 shows all of the processing schedules we used in our procedure. To simplify the experimental design, patterns were only included with three switches. Note that the amount of patterns with and without a consolidation opportunity was equal.

In both conditions, the length of the trials was progressive from 1 to 7 letters, with five trials per list length. The reaction times of the processing task were obtained by asking the participants to press on the space bar when they read the items aloud. Each trial started with a fixation cross displayed for 750 ms, followed by the letter that appeared for 1500 ms. After the consonant, the processing events (i.e. four digits and four slots of free time) were displayed, each for 600 ms. At the end of each trial, the participants were instructed to recall the letters in the correct serial order. The recall procedure was identical to Experiment 3 and 4. Before starting the experimental trials, a warm up was carried out similarly to our previous studies.

3.4.2 Results and discussion

Results from three participants were excluded from analysis as the proportion of correctly recalled items deviated more than two standard deviations from other participants. For the remaining 41 participants (mean age $= 22.0$ years, $SD = 4.4$), we used the same scoring method as in Experiments 2-4 (i.e. proportion of correctly recalled items in their correct order), referred as memory performance.

In light of the results of our previous experiment that showed a beneficial effect of an immediate delay of free time after the presentation of the storage event, we first wanted to verify, if there was a significant difference in memory performance between participants within the predictable condition. In this preliminary analysis,

1	L	0	0	0	1	0	1	1	1
2	L	0	0	0	1	1	1	0	1
3	L	0	0	1	1	0	0	1	1
4	L	0	1	0	0	0	1	1	1
5	L	0	1	1	0	0	0	1	1
6	L	1	0	1	1	1	0	0	0
7	L	1	0	0	1	1	1	0	0
8	L	1	1	0	0	1	1	0	0
9	L	1	1	1	0	0	0	1	0
10	L	1	1	1	0	1	0	0	0

Figure 3.10: Processing schedules of the concurrent task used in Experiment 5. *Note:* L = Letter displayed for 1500 ms. The duration of any processing event (i.e. free delay or digit) was 600ms. The dummy code 1 stands for a digit and 0 for a delay of free time.

we did not find a significant beneficial effect of consolidation on memory performance ($t(41) = -1.3, p = 0.2$). We then tested the main hypothesis, that is, the global effect of task predictability on memory performance. We used the same rational as in our previous experiments to build the statistical models. Table 3.12 summarizes these models.

Table 3.12: Summary of the statistical models created for Experiment 5

$M0 \longleftarrow$ DV$\sim$ + $(1|Subject)$
$M1 \longleftarrow$ DV$\sim$ List Length + $(1|Subject)$
$M2 \longleftarrow$ DV$\sim$ Predictability + $(1|Subject)$
$M3 \longleftarrow$ DV$\sim$ List length + Predictability + $(1|Subject)$
$M4 \longleftarrow$ DV$\sim$ List Length * Predictability + $(1|Subject)$

Table 3.13: Summary of model selection in Experiment 5 for memory performance based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M3$	5	648.11	0.00	0.48
$M4$	6	649.25	1.13	0.27
$M1$	4	649.50	1.39	0.24
$M2$	4	1483.41	835.30	0.00
$M0$	3	1483.93	835.82	0.00

Table 3.14: Mean memory performance (SD) as a function the processing patterns.

00010111	00011101	00110011	01000111	10011100
.65 (.20)	.64 (.20)	.66 (.18)	.64 (.22)	.60 (.22)

10111000	11000110	11001100	11100010	11101000
.63 (.23)	.60 (.23)	.63 (.22)	.64 (.22)	.62 (.20)

Table 3.13 shows the results of our model comparison. The model including list length and predictability of the processing events ($M3$) was ranked best. The evidence ratio between the first three models was low ($M3/M4 = 1.72$; $M3/M1 = 2$), which suggested equal support for these models. Regarding the parameters of the model best ranked, we observed significant differences in memory performance as a function of memory load ($F(1, 2828.4) = 975.11$, $p < .001$), with an estimate showing a -8% memory performance decrease when load was increased by 1. The factor predictability did not show any significant effect on memory performance

($F(1, 2829) = 3.4$, $p = 0.07$), with an estimate of -2% memory performance decrease when the concurrent task was unpredictable. Analysis of simple effects did not show a significant difference between the unpredictable condition and the predictable condition using pattern 00110011, neither for pattern 00110011. Table 3.14 shows the mean memory performance as a function of the processing patterns used in this experiment.

In conclusion, this experiment examined the potential effect of switching through task predictability. The idea was that switching costs in our previous experiments may have been neutralized by the participants' ability to prepare for a switch due to the familiarity of the processing patterns throughout their subsequent repetitions in longer list lengths. In order to test this hypothesis, we compared a predictable concurrent task using a repetitive simple pattern to an unpredictable one in which the processing events between the letters constantly varied. Our statistical analysis showed equal support for a model with or without the factor predictability. Our finding suggests thus that the predictability of the task has only a negligible effect on working memory performance. In line with our finding is the study by Bernardin, Portrat, and Barrouillet (2006) who examined the predictability of the concurrent task through the manipulation of the rhythm and the nature of the concurrent task. The authors compared in their study an unpredictable rhythm condition in which the duration of the free delay in between distractors varied from 200 ms to 1000 ms and a regular rhythm condition in which the interstimulus interval was set at 600 ms. The total duration of delays of free time was, however, identical in both conditions. Akin to our results, the authors observed that the rhythm of completion of the concurrent task had no influence on the linear relationship between cognitive load and memory performance. According to Bernardin et al. (2006), this observation suggests that maintenance and processing events in complex span task are managed by external cues which require minimal preparation of refreshing episodes.

In hindsight, it appears that our experimental design was not optimal to study this question. In this experiment, we did not manipulate the number of switches in the processing patterns in order to focus on whether the concurrent task was globally predictable versus not. However, perfect control of potential consolidation effects was lacking in our design[3]. In the predictable condition, participants were submitted to either a concurrent task with or without a consolidation opportunity.

[3]This might puzzle the reader since we have examined consolidation effects in our previous experiment. However, we conducted both Experiments 4 and 5 concomitantly. Also, it turns out that we could have been more cautious, if we had been aware of the study by Bayliss et al. (2015).

Although, we checked that no significant difference was apparent between the two patterns of the predictable condition, this null result could reflect a lack of statistical power. This issue is also valid for the two separate tests we carried out (each sub-group of the predictable task compared to the unpredictable task). In the unpredictable condition, the consolidation processing patterns (i.e. starting with a delay of free time) were equated with patterns that did not provide a consolidation opportunity. The main reason for this was that we believed that the unpredictability of the task would be driven by a minimal amount of patterns to be presented. With this idea in mind, we randomly presented a great number (i.e. 10) of different processing schedules across the unpredictable task.

However, the fact that potential consolidation processes in working memory studies are not yet fully understood questions the validity of our approach, in particular because the patterns in the unpredictable condition started with a consolidation opportunity higher (i.e. 1800 ms) than the one in the predictable condition (i.e. 1200 ms). This could have partially favored the unpredictable condition.

Concerns about the time course of consolidation processes and the nature of the relation between this process and memory or cognitive load are still debated (Souza & Vergauwe, 2018). It is arguable that the potential effect of consolidation is really neutralized by simply introducing processing schedules starting immediately with a distractor, especially when the patterns are randomly drawn for each list length. For example, if we assume that the consolidation patterns do enhance working memory capacity – provided that statistical power is sufficient – and that this effect is sensitive to storage demand (which is what our previous experiment suggests), then this could have easily introduced uncontrolled differences in the participants' performance.

Crucially, and this is the main reason why we will refrain to draw any strong conclusions regarding task-predictability effects in this experiment: all of the issues raised above make the comparison between the predictable and the unpredictable task uneasy. Since there is uncertainty about the exact effect of consolidation over time (although De Schrijver and Barrouillet (2017) suggest that the effect follows a decelerating function over time), it might be possible that our two experimental conditions might vary not only regarding our intended factor of task predictability but also regarding the consolidation factor.

In summary, the experimental designs used so far appear to present inherent limitations. As we have seen, possible confounds are easily introduced, and a better approach could be to focus on unambiguous processing patterns that could isolate the alleged factors of interest specifically.

3.5 Conclusion

Switching between tasks induces a toll on the cognitive system. Following the idea that complex span tasks require to switch between storage and processing activities and vice-versa, the present Chapter attempted to detect switching costs prompted by the structure of a task in working memory performance. In order to examine this question, we ran a series of experiments that introduced subtle variations in the processing schedules of complex span tasks. We argued that how storage and processing activities are distributed within a concurrent task should induce switching costs in recall and concurrent task performance.

Our first study examined switching costs by implementing a maximally randomized processing schedule in a complex span task. While this general design was interesting for the study of switching costs, it introduced a confound between cognitive load and the proportion of switches. This was, however, a starting point allowing us to test a large range of patterns. While our results did not bring credible evidence in favor of our hypothesis, it allowed to reveal that the most conducive condition to explore the effect of switching on performance is a medium cognitive load condition.

We, therefore, pursued our investigation by implementing switching through the use of predefined processing schedules, keeping the cognitive load constant. These experiments allowed us again to manipulate how storage and processing activities are distributed within the concurrent task. Closer scrutiny of our processing patterns enabled us to separate three different factors that could perhaps influence working memory performance. We assumed that switching and irregularity in the task should impair performance, whereas consolidation should benefit working memory performance.

It turned out that the processing task used in Experiment 3 did not allow to answer our research questions due to a ceiling effect. In Experiment 4 we, therefore, opted for a more demanding concurrent task, while still keeping the cognitive load constant. The results of this study confirmed our hypothesis and showed that how processing events are organized within a task do influence working memory performance. Specifically, our result showed a beneficial effect of a free delay presented immediately after the storage event, which could be interpreted as evidence in favor of consolidation. However, the effect of this free delay on the concurrent task was rather elusive. We discussed alternative explanations to account for the consolidation effect in terms of task preparation and predictability, and also in terms of the refreshing schedules. Chapter 5 will focus in more depth on this latter alternative account of consolidation. Central to the present thesis, evidence in favor of switching costs on memory was not observed. In contrast, although this was not hypothesized, the results on concurrent task performance showed a beneficial effect when the processing schedule involved numerous alternations. We discussed the possibility to increase the potential effect of switching by increasing the cognitive load of the task. This aspect will be explored in Chapter 4. Finally, the results of this study did not support the idea that the irregularity of the processing patterns (i.e. the fact that global patterns in the task could be recognized and thus anticipated) could influence working memory performance.

The last experiment of this Chapter attempted to examine the effect of switching through task predictability and the use of irregular and regular processing schedules. Our result did not show evidence in favor of a detrimental effect of task unpredictability, although we pointed out several methodological issues related to the inherent limitation of our experimental material. We concluded that a better approach would be to focus on a more restricted amount of processing patterns that could more easily isolate the alleged factors of interest specifically.

Overall, the studies of this Chapter revealed a more complex reality than what we had initially envisioned. The main finding here is that processing schedule variations might influence memory performance in some cases but not in others. This points out to the complexity of the object of study, as several confounds are easily introduced. In an effort to reduce these possible confounds, in the following Chapters we will focus on a limited amount of processing schedules, that were designed to explore the switching cost and the consolidation hypotheses.

4 Switching in complex span tasks

In our previous Chapter, we established that the manipulation of processing schedules may influence working memory performance in some cases. However, our results did not bring credible evidence in favor of the idea that switching could cause the observed variations of performance. The following experiments presented in this Chapter were designed to test whether a higher cognitive load of the task at hand could better reveal switching costs in complex span tasks. The rationale is that the cost involved in task switching might depend on the task difficulty. The idea is that a more challenging task might require more time to load and unload its task set compared to a simpler task.

The studies presented in this Chapter made use of similar processing patterns than those used in Experiment 3 and 4. We focused specifically on two patterns manipulating the number of switches between storage and processing activities. These chosen patterns did not start with a delay of free time, nor they were irregular in terms of the number of distractors they presented in a row. This should ease the interpretation of the root cause of possible variations of task performance. Furthermore, the cognitive load of the task was manipulated across blocks of conditions, and not across trials to avoid the caveats of Experiment 2. Increased cognitive load in the complex span task was achieved by augmenting the attentional capture and the pace of the concurrent task. Such a manipulation of the cognitive load allowed to keep the number of distractors constant in order to avoid a confound between cognitive load and the amount of switches in a given trial. We expected to find reduced working memory performance as a result of a high number of switches to execute between storage and processing activities in conditions where the cognitive load of the task at hand was high.

4.1 Experiment 6a

4.1.1 Method

Participants. One hundred and thirty students of Université Côte d'Azur participated to this experiment in exchange for partial course credit (107 females, 22 males; mean age = 21.68 years $SD = 4.9$). All participants were tested in a quiet room at the university. The duration of the experiment was approximately one hour.

Material and procedure. In this experiment, participants were invited to memorize letters while completing a parity judgment task. Participants were randomly assigned between two experimental groups differing in the pace of the concurrent task (1200 ms *vs.* 900 ms per processing event). Within these two experimental groups all participants were submitted to two conditions randomly ordered for each participant in which the parity judgment task was executed on digits in their Arabic or Roman forms. Previous research has shown that processing roman numerals is more time consuming than arabic numerals (e.g. Barrouillet et al., 2007). Following this finding, our manipulation of the form of digits was thus meant to vary the attentional capture of the concurrent task. Taken together, the manipulation of the pace and the attentional capture of the task allowed us to create four levels of cognitive load: Fast-High (Roman numerals displayed at 900 ms.) *vs.* Slow-High (Roman numerals displayed at 1200 ms) *vs.* Fast-Low (Arabic numerals displayed at 900 ms) *vs.* Slow-Low (Arabic numerals displayed at 1200 ms). Each participant completed the concurrent task at either a fast pace or at a slow pace.

The material of the storage items were the following consonants: B, F, H, J, K, L, N P, Q, R, S, T. The processing items were digits randomly drawn from 1 to 9 in their Arabic or Roman forms, depending on the condition. In both these conditions, the patterns of the processing task were either composed of three switches or seven switches. Each pattern was repeated two times per set size in a random order. The processing patterns are shown in Figure 4.1. As seen, both patterns started immediately with a digit in order to avoid any confounds with a consolidation effect.

Each trial started with a fixation cross displayed at the center of the computer screen for 750 ms, followed by a letter presented for 1500 ms. After the storage event, a digit was presented for 1200 ms and the participant was invited to categorize the digit as odd or even by pressing the letters I or P respectively on the keyboard as quickly as possible. A response was considered incorrect for the parity judgment task whenever the participant failed to answer in the allowed time window. Reaction times were recorded from the presses of the keyboard. The set size of the trials progressively increased in both conditions from 3 letters to 7 letters. Each set size was repeated four times.

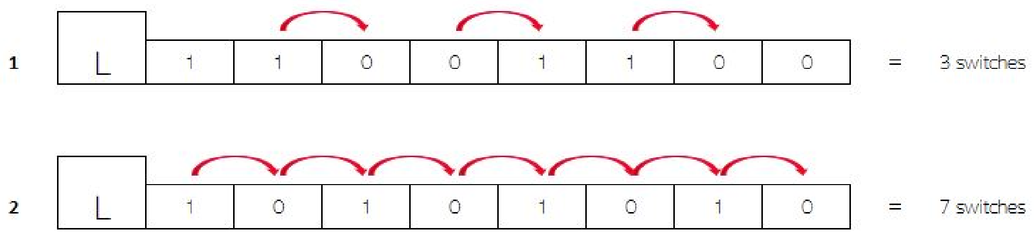


Figure 4.1: Processing patterns of the concurrent task used in Experiment 6a and 6b. *Note:* L stands for a letter (presented for 1500 ms), 1 stands for a digit and 0 for a delay of free time.

4.2 Experiment 6b

In the following, we will describe the design of experiment 6b, after which we will report the results of both experiments 6a and 6b together.

4.2.1 Method

Participants. One hundred thirty-four students of Université Côte d'Azur participated to this experiment in exchange for partial course credit (126 females, 8 males; mean age $= 21.21$ years, $SD = 2.0$). All participants were tested in a quiet room at the university. The duration of the experiment was approximately one hour.

Material and procedure. The material and procedure was identical to Experiment 6a. The only difference with the previous experiment was that participants were required to give oral responses to the parity judgment task while pressing on the space bar of the keyboard. This design was thought to enable the measure of the

RTs corresponding to the oral response to the parity judgment task. For instance, if the digit 7 (or VII in the roman numeral condition) was displayed, the participant was expected to say seven - odd in french (i.e. "sept - impair" and to press on the space bar at the same time.

4.2.2 Results of Experiments 6a and 6b

In both experiment, twelve participants in total were excluded as their performance at the concurrent task or recall task deviated more than two standard deviations from other participants. In experiment 6a, for the remaining one hundred twenty-nine participants, the mean age was 21.2 years ($SD = 2.0$). In experiment 6b for the remaining one hundred twenty-two participants the mean age was 21.8 years ($SD = 5.1$). We used the same scoring method as in our previous experiments (i.e. proportion of correctly recalled items in their correct serial position), referred to as memory performance from now on.

Effect of switching on memory performance. For both experiments 6a and 6b, we conducted a linear mixed model effect analysis to see if the number of switches between storage and processing activities in the processing schedules had an influence on memory performance. To this end, we proceeded to model comparisons as in Chapter 3. The statistical models and their ranking are presented in Table 4.1 and Table 4.2. Ranking of the models is based on their fit to the data, which was computed using the number of parameters of each model (K), the AICc values, the delta AICc values and the weight of each model for each experiment.

Table 4.1: Statistical models created for Experiment 6a and 6b.

$M0 \longleftarrow$ DV∼ + $(1|Subject)$
$M1 \longleftarrow$ DV∼ CL + $(1|Subject)$
$M2 \longleftarrow$ DV∼ Number of switches + $(1|Subject)$
$M3 \longleftarrow$ DV∼ CL + Number of switches + $(1|Subject)$
$M4 \longleftarrow$ DV∼ CL * Number of switches + $(1|Subject)$

For experiment 6a, when the concurrent task involved non-verbal responses, the results showed that the model incorporating only cognitive load ($M1$) was best ranked, with a strong evidence ratio against the null model ($M1/M0 > 150$). In contrast, and against our hypothesis, the model incorporating only the number of switches was ranked last, meaning that the influence of switching on memory performance was null. Regarding the parameters of interest of the model best ranked, we observed an effect of cognitive load ($F(3, 368.66) = 9.28$, $p < .001$). In reference to the Fast-High cognitive load condition, the estimates showed a memory performance increase of 3% and 2.3% for the Fast-Low and Slow-Low cognitive load conditions respectively, and a decrease of -1.2% for the Slow-High condition. However, only the Fast-High and Fast-Low cognitive load conditions revealed significant differences on memory performance ($t(128) = 3.4$, $p < .001$; all other $ps > .21$). Figure 4.2 shows the mean memory performance as a function of cognitive load. As seen, all four cognitive load conditions revealed high memory performance.

Table 4.2: Summary of model selection in Experiment 6a and 6b for memory performance based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 4.1.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
		Experiment 6a		
$M1$	6	-560.44	0.00	0.67
$M3$	7	-558.48	1.96	0.25
$M4$	10	-556.23	4.21	0.08
$M0$	3	-538.70	21.74	0.00
$M2$	4	-536.74	23.70	0.00
		Experiment 6b		
$M3$	7	2328.29	0.00	0.91
$M4$	10	2332.81	4.51	0.09
$M1$	6	2345.11	16.82	0.00
$M2$	4	2567.08	238.79	0.00
$M0$	3	2583.85	255.56	0.00

Interestingly, when the concurrent task involved verbal responses in experiment 6b, the model best ranked was the one including cognitive load and the number of switches ($M3$). When both factors were compared to the null model separately, the results indicated strong evidence that both factors contributed to memory performance ($M1/M0 > 150$; $M2/M0 > 150$). Finally, we found equal support for the full models with and without the interaction term ($M3/M4 = 0.1$).

Regarding the parameters of interest of the model best ranked, we observed an effect of cognitive load ($F(3, 396.6) = 83.823$, $p < .001$). In reference to the Fast-High cognitive load condition, the estimates showed a memory performance increase of 13%, 16% and 23% for the Slow-High, Fast-Low and the Slow-Low cognitive load conditions respectively. All pairwise comparisons revealed significant differences on memory performance. In sum, the manipulation of cognitive load in this experiment was more effective than in our previous study.

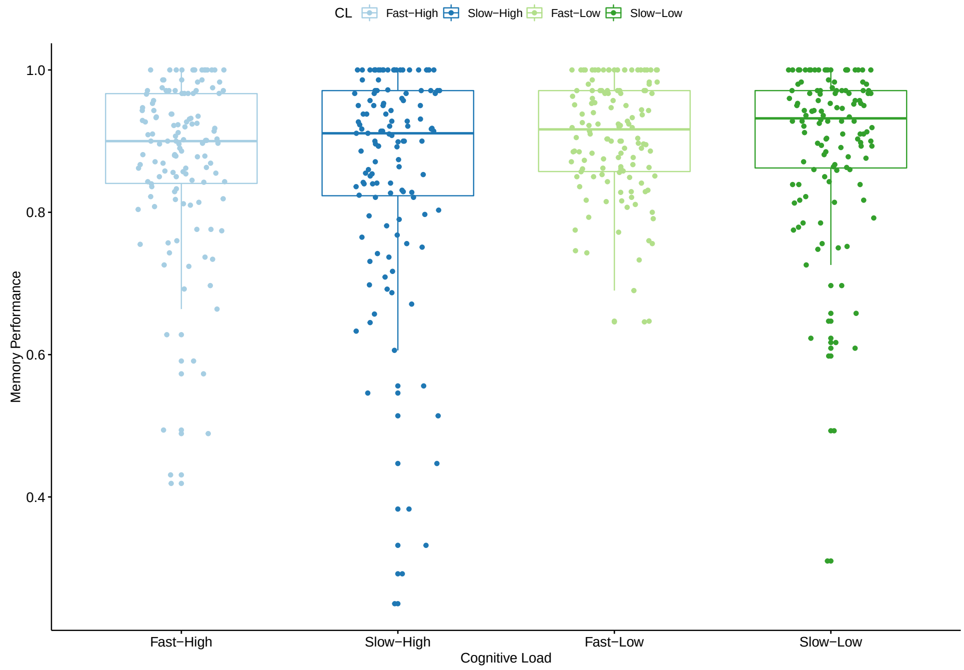


Figure 4.2: Mean memory performance as a function of cognitive load in Experiment 6a. *Note:* Cognitive Load (CL) = Fast-High (Roman numerals displayed at 900 ms.) *vs.* Slow-High (Roman numerals displayed at 1200 ms) *vs.* Fast-Low (Arabic numerals displayed at 900 ms) *vs.* Slow-Low (Arabic numerals displayed at 1200 ms)

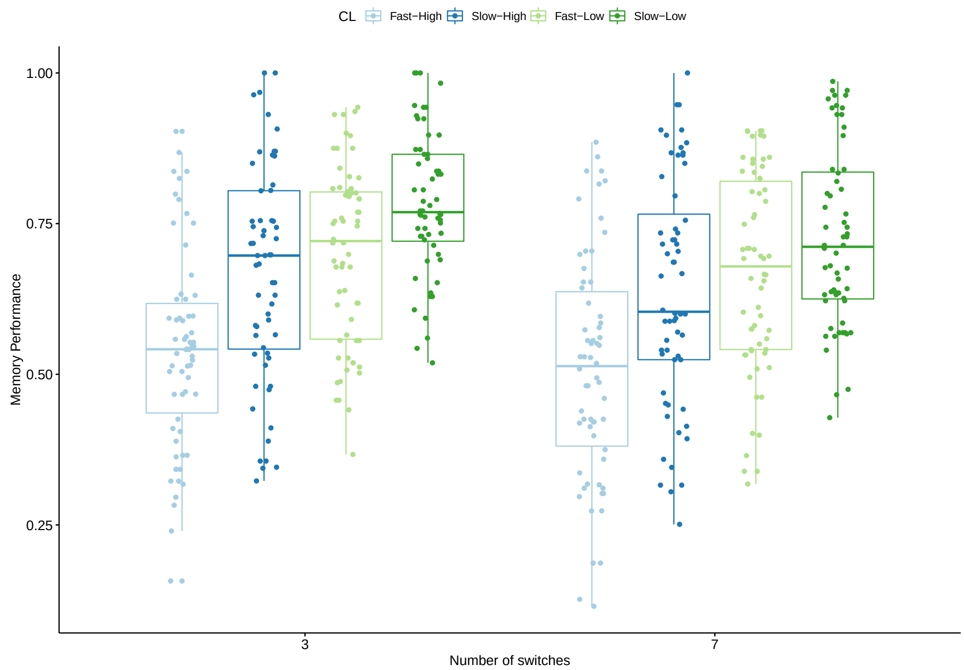


Figure 4.3: Mean memory performance as a function of cognitive load and the number of switches in Experiment 6b.

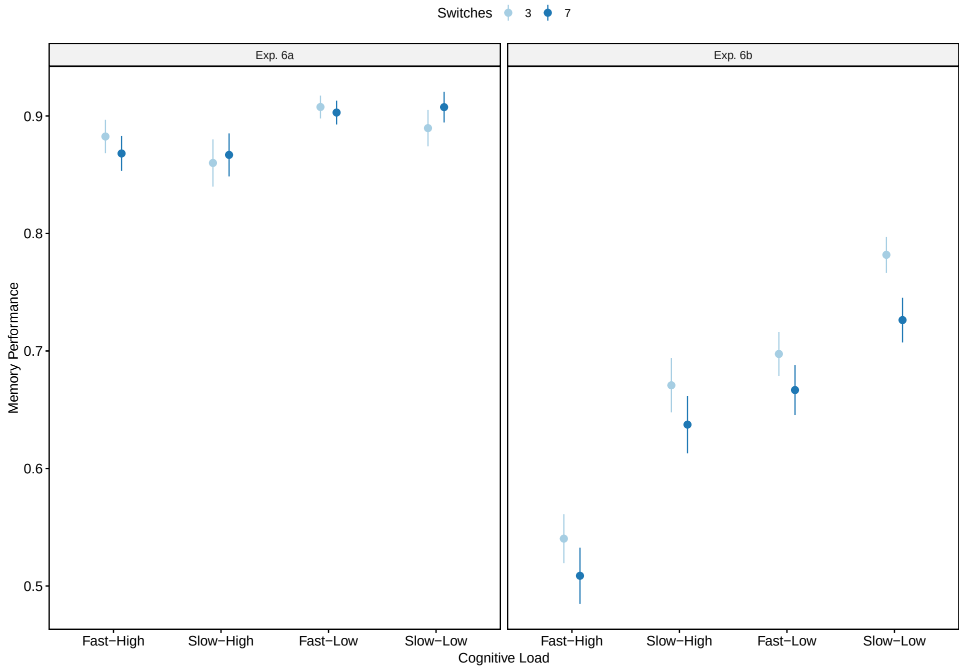


Figure 4.4: Mean memory performance as a function of cognitive load and the number of switches in Experiment 6a and 6b. *Note:* Cognitive Load (CL) = Fast-High (Roman numerals displayed at 900 ms.) *vs.* Slow-High (Roman numerals displayed at 1200 ms) *vs.* Fast-Low (Arabic numerals displayed at 900 ms) *vs.* Slow-Low (Arabic numerals displayed at 1200 ms).

Regarding our second variable of interest, results showed an effect of the number of switches on memory performance ($F(1, 4761.1) = 18.862$, $p < .001$). The estimates showed a memory performance decrease of 4% when the number of switches in the processing schedule increased from three to seven. Figure 4.3 shows the mean memory performance as a function of cognitive load and the number of switches. As seen, although significant differences in memory performance were found in our analysis, the effect of switching still appears to be small.

Figure 4.4 summarizes our results and shows mean memory performance in experiments 6a and 6b as a function of cognitive load and the number of switches. Again, memory performance in experiment 6a did not vary substantially as a function of cognitive load, and the switch effect was null. With regards to experiment 6b, we saw a clear detrimental effect of cognitive load on memory performance. Additionally, we saw a clear tendency for better performance when the processing task involved a fewer switches between storage and processing activities. This benefit was most apparent in the Slow-Low cognitive load condition in which the pace of the secondary task was 1200 ms and the numerals were presented in their Arabic form.

Effect of switching on RTs. For both experiments 6a and 6b, we conducted the same analysis as in the previous section to examine the effect of switches on reaction times in the concurrent task (see Table 4.1 for a summary of the statistical models). Our dependent variable was the mean Reaction Time (RT) in ms for the parity judgment operations on a trial-by-trial basis (incorrect RTs were excluded from analysis).

Table 4.3 shows the ranking of the models for each experiment when the dependent variable was the mean RT. For experiment 6a, when the processing task involved non-verbal responses, the results showed that the full model with the interaction term ($M4$) was best ranked. The calculated evidence ratio of both models including only one factor of interest was strong ($M1/M0 > 150$; $M2/M0 > 150$), which suggests that both cognitive load and the number of switches contribute together to the RTs of the processing task. Finally, the calculated evidence ratios between the full models with and without the interaction term indicated clear support for the interaction model, but the simpler model could not yet be rejected ($M4/M3 = 42.17$).

Regarding the parameter of interest of the model best ranked in Experiment 6a, both main effects on RTs were significant ($F(3, 405.6) = 395.89$, $p < .001$; $F(1, 5010.0) = 62.77$, $p < .001$ for the factor cognitive load and the number of switches respectively). Additionally, the interaction between cognitive load and the number of switches was also significant ($F(3, 5010.0) = 4.51$, $p < .001$). With regards to the Fast-High CL condition, the estimate showed an increase of RTs of 6 ms in the Slow-High CL condition, and a decrease of 28 ms and 54 ms for the Slow-Low and Fast-Low conditions respectively. Regarding the switching estimate, it showed an increase of RTs of 16 ms when the number of switches increased from three to seven. Figure 4.5 shows the mean RTs of the parity judgment operations as a function of cognitive load and the number of switches. As seen, the effect of the number of switches is clear for both Fast-High and Medium High conditions, but not for Slow-Low condition.

Experiment 6b yielded essentially the same results as Experiment 6a, while confirming that the full model including the interaction term was superior to the simpler model ($M4/M3 > 150$). Regarding the parameter of interest of the model best ranked in Experiment 6b, results showed an effect of cognitive load and the number of switches on RTs ($F(3, 349.4) = 994.12$, $p < .001$; $F(1, 4758.4) = 279.94$, $p < .001$ for the cognitive load factor and the number of switches respectively).

Table 4.3: Summary of model selection in Experiment 6a and 6b for RT data based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 4.1.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
		Experiment 6a		
$M4$	10	55595.57	0.00	0.98
$M3$	7	55603.06	7.48	0.02
$M1$	6	55663.20	67.63	0.00
$M2$	4	56662.12	1066.55	0.00
$M0$	3	56710.53	1114.96	0.00
		Experiment 6b		
$M4$	10	57091.93	0.00	1
$M3$	7	57121.83	29.90	0.00
$M1$	6	57390.97	299.04	0.00
$M2$	4	59431.81	2339.88	0.00
$M0$	3	59590.29	2498.36	0.00

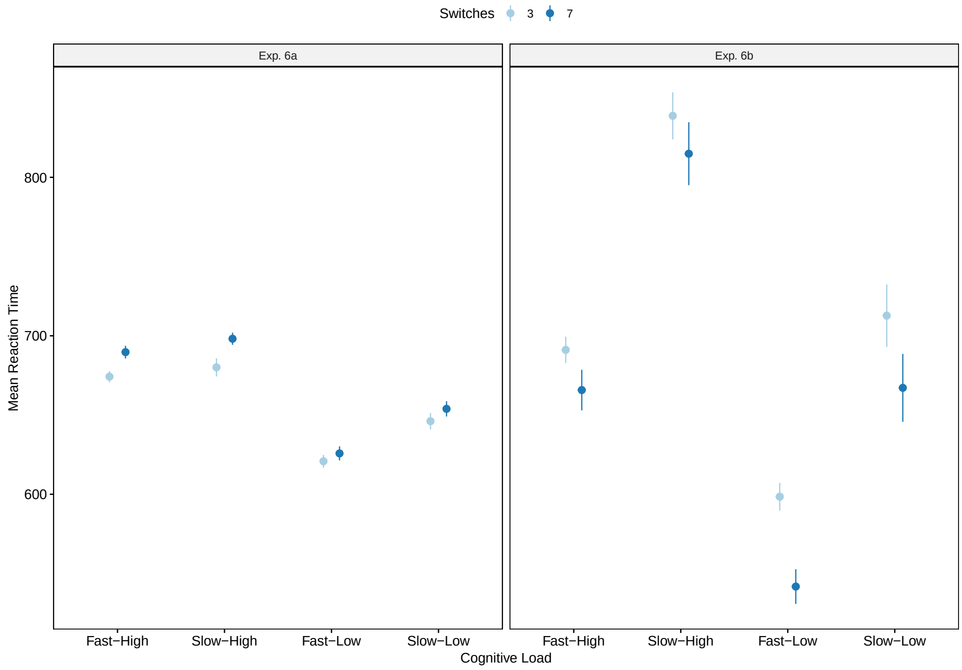


Figure 4.5: Mean RT of the concurrent task as a function of cognitive load and the number of switches in Experiment 6a and 6b. *Note:* Cognitive Load (CL) = Fast-High (Roman numerals displayed at 900 ms.) *vs.* Slow-High (Roman numerals displayed at 1200 ms) *vs.* Fast-Low (Arabic numerals displayed at 900 ms) *vs.* Slow-Low (Arabic numerals displayed at 1200 ms).

Additionally, the interaction between cognitive load and the number of switches was also significant ($F(3, 4758.4) = 12.02$, $p < .001$). Regarding the estimate, and in reference to the Fast-High CL condition, results showed an increase of 150 ms and 23 ms of RTs in the Slow-High and the Slow-Low condition, and a decrease of 92 ms in the Fast-Low condition. Interestingly, the estimate for switching in this experiment showed the opposite trend compared to Experiment 6a, with a decrease of RTs of 24 ms when the concurrent task varied from three to seven switches, as seen from Figure 4.5.

Effect of switching on concurrent task performance. In the following, we report the analysis of the concurrent task of experiment 6b. As our dependent variable we used the proportion of correctly answered parity judgment operations. The same statistical models were used as in our previous analysis (see Table 4.1 for a summary.

Table 4.4 shows a ranking of the models in terms of of the best fit for each experiment. The results showed that the model incorporating cognitive load and the number of switches without the interaction term ($M3$) was ranked best when the dependent variable was the performance at the concurrent task. When both factors were compared to the null model separately, the results indicated strong evidence that both factors contributed to performance at the concurrent task ($M1/M0 > 150$; $M2/M0 > 150$). Regarding the comparison between the model with and without the interaction term, evidence pointed out to equal support between the two models ($M3/M4 = 1.84$).

Table 4.4: Summary of model selection in Experiment 6b for concurrent task performance based on the following values: number of parameters of each model (K), AICc values, delta AICc values and weight of each model described in Table 4.1.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M3$	7	-6805.91	0.00	0.65
$M4$	10	-6804.69	1.21	0.35
$M1$	6	-6791.12	14.79	0.00
$M2$	4	-5346.13	1459.78	0.00
$M0$	3	-5335.58	1470.33	0.00

Regarding the parameters of interest of the model best ranked, we observed an effect of cognitive load ($F(3, 375.2) = 566.61$, $p < .001$). In reference to the Fast-High cognitive load condition, the estimates showed a concurrent task performance increase of 7.5%, 16% and 19% for the Slow-High, Fast-Low and the Slow-Low cognitive load conditions respectively. Pairwise comparisons of the different cognitive load conditions all revealed significant differences on the concurrent task performance. Additionally, results showed an effect of the number of switches on the processing task performance ($F(1, 5027.0) = 16.82$, $p < .001$). As the number of switches increased from three to seven switches, the estimate indicated a small increase of 1.4% of the concurrent task performance.

Figure 4.6, shows the mean concurrent task performance as a function of cognitive load and the number of switches. Although, the manipulation of cognitive load in Experiment 6a had a less apparent effect on memory performance as we discussed earlier, the effect on the concurrent task performance is clear as seen in the figure. As to the effect of switching, a small trend is visible on visual inspection with better performance when a higher amount of switches needs to be performed.

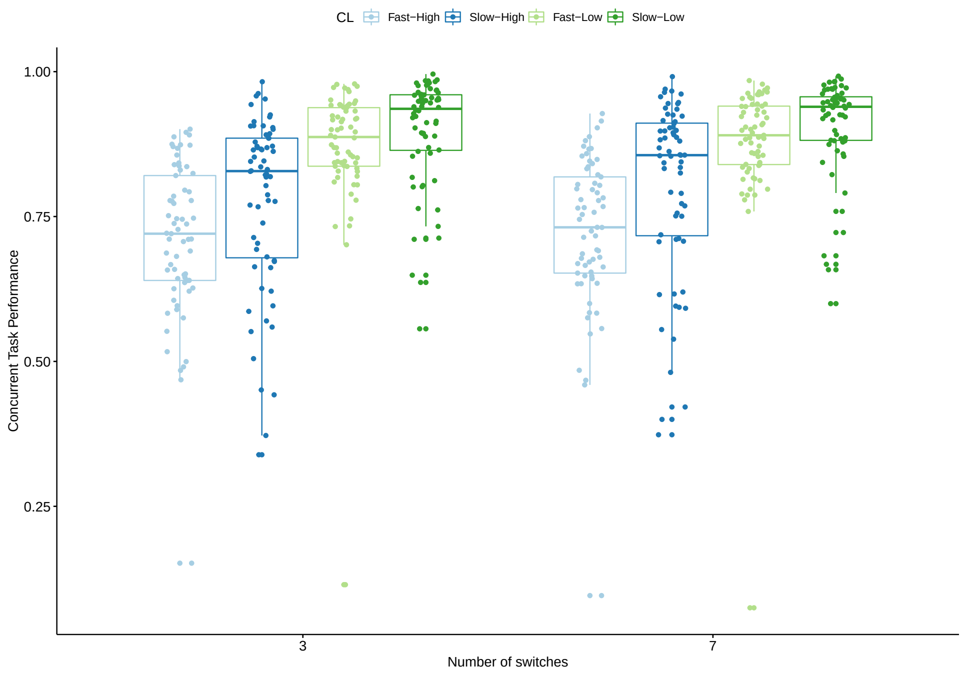


Figure 4.6: Mean performance at the concurrent task as a function of cognitive load in Experiment 6a. *Note:* Cognitive Load (CL) = Fast-High (Roman numerals displayed at 900 ms) *vs.* Slow-High (Roman numerals displayed at 1200 ms) *vs.* Fast-Low (Arabic numerals displayed at 900 ms) *vs.* Slow-Low (Arabic numerals displayed at 1200 ms).

4.2.3 Discussion

The present study examined the effect of switching between storage and processing in complex span tasks by varying the processing schedules and cognitive load.

We expected to find reduced working memory performance as a result of a high number of switches between storage and processing activities in conditions where the cognitive load of the task at hand is high.

Our results showed only partial confirmation of our hypothesis. In detail, we found no evidence in favor of switching costs in memory performance regardless of the cognitive load of the task at hand when no verbalization was required (Experiment 6a). Regarding the concurrent task for this same experiment, the results showed a beneficial effect of the switches on accuracy, meaning that when participants were required to switch more often they performed better. This was, however, achieved at a cost, since participants were also slower when they performed better. This finding suggests that accuracy-speed adjustments were made by the participants. When the complex span task involved a parity judgment task requiring verbalization (Experiment 6b), we observed evidence in favor of small switch costs on memory performance, in particular for the condition involving a low cognitive load (induced by an easy concurrent task) and a slow pace, although we had predicted that the switching costs would be most apparent in the high cognitive load conditions. In contrast, the response times of the concurrent task showed a beneficial effect of switching when the number of switches to execute was high. In other words, faster responses were associated to a higher number of switches. Unfortunately, we did not collect the accuracy data of the concurrent task for this experiment, so we are not able to know whether participants made in this case speed-accuracy adjustments during the processing task which could perhaps explain this result. Nevertheless, the fact that we found a switch cost for the processing speed of the concurrent task (Exp. 6a), and a switch cost for the memory task (Exp. 6b) is a novel finding, since previous work had not identified a detrimental effect of switching between storage and processing activities in a complex span task.

Overall, the findings of these experiments seem to suggest that switching costs in complex span task are sensitive to its specific design. The fact that switching appeared to have a detrimental effect on recall performance in our second experiment, but not in the first one, is in line with previous work that has suggested that verbalization plays an important role in task-switching behavior (e.g. Bryck & Mayr, 2005; Liefooghe et al., 2005; Miyake et al., 2004; Saeki & Saito, 2004, 2009). However, the sole role of verbalization do not allow to explain our results, since it appears that our manipulation of switching was the most effective in the lowest and slowest cognitive load condition, suggesting that the attentional demand of the task is crucial to detect switch costs in memory performance. This finding could suggest that when the task is too demanding, participants might not have enough time or available attentionnal resource to switch back and forth. In contrast, when the cognitive load of the task is too low, the switch might not tax performance visibly; despite the switch cost participants could be able to compensate for the cost and reach their span level because they have sufficient time remaining to do so. This might also explain the lack of positive findings on recall performance in Experiment 6a. The manipulation of the cognitive load produced in this case only little variation of memory performance, and all four conditions showed relatively high memory performance, suggesting that the task was not demanding enough.

Closely related to the idea that switch costs are dependent of an optimal cognitive load condition is the notion of task prioritization. Participants might choose to prioritize one or the other task. For instance, they might prioritize the concurrent task performance over the memorization or vice versa (Belletier, Camos, & Barrouillet, 2020; Rhodes et al., 2019). Such a strategic choice would mean that participants do not switch regularly between storage and processing activities. Furthermore, it is also likely that participants might make strategic speed-accuracy adjustment in highly demanding tasks within the concurrent task (Draheim, Hicks, & Engle, 2016), as we observed in Experiment 6a. Finally, it is also possible that these prioritization vary during the execution of the task, with participants favoring at some point different aspects of the task at hand, making thus the detection of switch costs difficult. This could also explain in part why switch costs appear small, and why they are not systematic from one study to another, or even from one cognitive load condition to another in the present experiments.

4.3 Conclusion

The goal of the present Chapter was to continue our investigation on the effect of switching between storage and processing on working memory performance. We presented here two experiments which manipulated switching and cognitive load. We assumed that switching costs would be most apparent in high cognitive load condition.

In both experiments we used a parity judgment task. The first experiment required from the participant keyboard responses to the concurrent task, whereas our second experiment required oral responses. The results yielded novel findings, with a partial confirmation of our hypothesis.

In short, the complex span task requiring verbal responses showed effectively small switching costs associated to the recall performance, whereas the task requiring keyboard responses did not. Regarding the concurrent task performance, we detected switching costs associated to the complex span task requiring keyboard response, whereas the complex span task requiring verbal response showed enhanced performance when a high number switches were to be executed. We discussed the discrepancies between the two studies in terms of the potential role of verbalization, cognitive load, and task prioritization. Further investigation is needed to disentangle which factors could contribute, if not all, to our findings.

5

Consolidation and refreshing in complex span tasks

The concept of consolidation was initially introduced in the long-term memory literature over a 100 years ago (Lechner, Squire, & Byrne, 1999). Because principles of parsimony call for a unitary view (Surprenant & Neath, 2009), the concept has found its way to apply to working memory processes as well. Research on consolidation within the working memory paradigm has thus dramatically increased, and this only in a few years (e.g. Blalock, 2015; Nieuwenstein & Wyble, 2014; Ricker & Hardman, 2017; Vergauwe, Ricker, Langerock, & Cowan, 2019). Nonetheless, consolidation as a specific working memory maintenance mechanism remains a controversial topic. As established by several reviews (C. C. Morey & Cowan, 2018; Ricker, Nieuwenstein, Bayliss, & Barrouillet, 2018; Souza & Vergauwe, 2018), this lack of consensus is mostly due to the fact that the concept itself is poorly defined. More importantly, in regard to the present study, it is uncertain *whether* and *how* consolidation is distinct from other already known working memory maintenance processes, such as refreshing (e.g.Camos et al., 2009; Raye, Johnson, Mitchell, Greene, & Johnson, 2007; Souza, Rerko, & Oberauer, 2015). Effectively, the behavioral predictions of the two processes can be surprisingly similar.

Consolidation in working memory and refreshing are both thought to be attention-based processes that contribute to the maintenance of memory traces, but very little is known about how they relate to one another and whether they are truly separate processes. By definition, consolidation stabilizes and strengthens novel information (Jolicœur & Dell'Acqua, 1998). Refreshing, on the other hand, reactivates already existing memory traces to avoid their loss (Camos et al., 2018). According to C. C. Morey and Cowan (2018), one hypothesis is that consolidation and refreshing could operate sequentially. According to this view, consolidation would occur immediately after the early stages of sensory and perceptual encoding, whereas refreshing would occur later on, once information is consolidated. However, it remains unclear when exactly consolidation would end and refreshing begin. According to recent findings, consolidation occurs between 400 ms and 5000 ms (e.g. De Schrijver & Barrouillet, 2017; Nieuwenstein & Wyble, 2014; Ricker & Hardman, 2017; Stevanovski & Jolicœur, 2007). Other more debatable account reports rates as fast as 50 ms per item (Vogel, Woodman, & Luck, 2006). The time course of refreshing appears to be faster, and it is estimated to last between 40 ms and 80 ms per item[1] (e.g. Jarrold et al., 2011; Oberauer & Lewandowsky, 2011; Vergauwe et al., 2014; Vergauwe & Cowan, 2014, 2015).

In working memory experiments, consolidation and refreshing are often manipulated by varying the amount of free delay between storage and processing events. The crucial difference between consolidation and refreshing resides mainly in *when* this free delay is presented (C. C. Morey & Cowan, 2018). For instance, in their complex span tasks, Bayliss et al. (2015) manipulated the position of the first processing item, which was presented either immediately after the storage item or delayed by 1000 ms respectively in their Experiment 2. In the immediate condition, this extra free time of 1000 ms was postponed to the end of the to-be-recalled sequence to keep the total duration of the task constant. The results showed that providing a consolidation opportunity to the participants (i.e. a delayed presentation of the first processing event) produced better memory performance. Further evidence was provided by De Schrijver and Barrouillet (2017), who observed higher working memory spans associated with longer consolidation periods. These findings were again interpreted as evidence in favor of a consolidation process, although the authors could not rule out the possibility of a refreshing account.

The reason delayed presentation of concurrent events produces better performance in working memory is not fully understood. Ricker and Cowan (2014) invoke several hypotheses to explain this phenomenon. One supposition is that free time after presentation of a to-be-stored item is used to implement more efficient strategies to prevent decay. However, this account does not clarify how exactly the process of consolidation would differ from maintenance processes. A second account of consolidation comes from a neural perspective. According to this view, the synchronization of firing neurons would create a short-term memory trace and their desynchronization would cause forgetting. Within this framework, consolidation is viewed as a gradual strengthening of this synchronization process. Ricker and Cowan (2014) proposed a functional explanation of this process based on the idea that the synchronization of firing neurons would only occur when attention is focused on the initially formed memory trace. The next section presents an alternative account of consolidation based on the process of refreshing.

[1] However, note that recent work evokes the possibility of two distinct refreshing processes : a high-speed automatic refreshing process and a more deliberate strategic, thus perhaps slower, refreshing process (Camos et al., 2018).

5.1 TBRS(2) & consolidation

Recall that the TBRS model (Barrouillet et al., 2004, 2007; Barrouillet & Camos, 2009) argues in favor of a time-related decay of memory traces in short-term memory, based on four key assumptions: (1) storage and processing in working memory rely both on attention, (2) focus of attention is directed either only on storage or on processing activities, (3) the activation of a to-be-stored item increases when attention is focused on it but decreases otherwise, and (4) what precedes implies that attention is required to constantly switch between storage and processing activities. Finally, the concept of cognitive load stemming from these premises is defined as the proportion of time during which a given task occupies attention. The next section describes how cognitive load can be mathematically approached and how we believe it can help decipher how consolidation works.

Recall that aside the putative switch cost, the TBRS2 model includes 4 parameters:

- The *decay rate* d corresponds to how fast the memory traces decay.
- The *refreshing rate* corresponds to how fast a memory trace is reinforced.
- The *duration* is the time devoted to refreshing a memory trace.
- The *baseline*

Following the theoretical logic of the TBRS, it should not matter how refreshment is distributed across items. Figure 5.1 (top vs bottom) illustrates this phenomenon to show the predictions of the TBRS2 when the opportunity for consolidation is manipulated while the cognitive load remains constant. The top of each plot displays the storage and processing events as a function of time. Each color represents a different item. All the black events represent the secondary task. The line 'Focus' represents the items to-be-remembered that must be maintained successively at the center of attention for later recall. As seen, free time allows the different memory traces to be refreshed alternatively. The curves below show the level of activation of each storage item. After 12 seconds (at the far right side of the frame), the level of activation predicts the probability of correct recall for each item. As shown by the numbers that are reported in the figure's note, the probability of correct recall is exactly the same *without* a consolidation opportunity (i.e. in the plot at the top, when processing events are placed immediately after the items) or *with* a consolidation opportunity (i.e. plot at the bottom with the presence of spare time immediately after the items are presented). The activation patterns along the time line however slightly differ from one condition to another, as explained below.

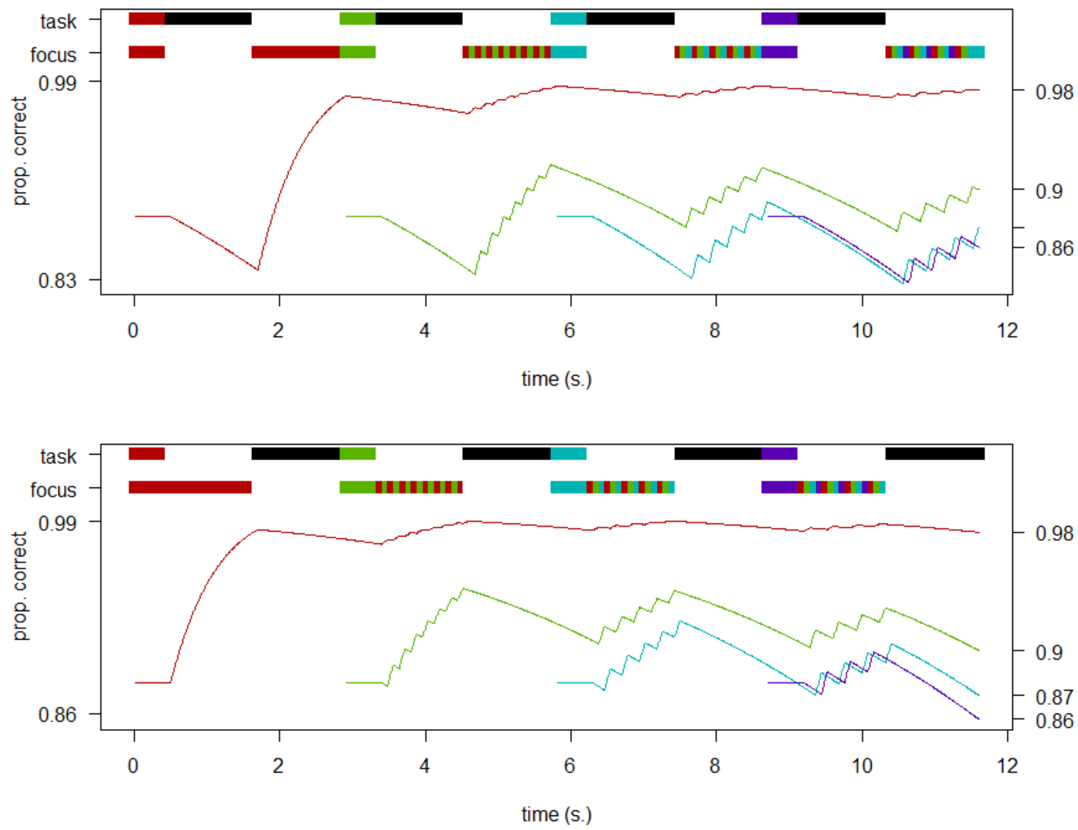


Figure 5.1: Top: Predictions of the TBRS2 for a sequence of four letters each immediately followed by a processing episode. Bottom: Predictions of the TBRS2 for a sequence of four letters starting with a consolidation opportunity. Both simulations involved a sequence of four letters and the following parameters: Duration = 80ms, Baseline = 5, $d = 1.4$, $r = 9$, event speed = 60, letter speed = 50. The probability of correct recall for the successive four letters are respectively 0.999, 0.999, 0.994, 0.984, for both cases.

Recall that, in the original TBRS model, attention can only be devoted to a single task and the only predictor of memory performance is cognitive load alone (i.e. the only factor influencing the activation of a memory trace is the amount of time devoted to refreshment). Once the mathematical constraints implemented in the TBRS2 model to mimic a pure effect of cognitive load, it is difficult to assume that the probability of recall of a set of items only depends of the cognitive load irrespective of the time schedule. Imagine a task in which one has to recall one item after an hour, during which attention is directed away every two seconds (for one second). Under the cognitive load assumption, the probability of recall should be the same for another task with the same cognitive load of 50%. If true, the TBRS should predict a similar recall performance of the same item even though attention is directed away continuously for the first 30 minutes, followed by a 30-minute period of free time before recall. However, the model has proven fairly powerful so far to account for empirical results in plausible working memory tasks in which immediate recall is expected, thus limiting potential paradox with longer delays (Barrouillet et al., 2011; Barrouillet & Camos, 2012).

Finally, note that the model can be implemented with exponential functions predicting the odds of recall of the items (not probabilities; see Gauvrit and Mathy (2018)). The odds can adequately represent correct recall as a function of the cognitive load, and thus allow a straightforward implementation of the model. Simply put, although the product of the probabilities of recall can not strictly be tied to cognitive load (because no mathematical function is mathematically satisfying), the product of odds is.

The difficulty to implement the theoretical model with probability values is of utmost importance when it comes to the question of the putative consolidation effect. Indeed, some empirical data show that in some circumstances at least, a drop-down of the span can be observed at constant cognitive load when the so-called consolidation (i.e. a refreshment period is present right after the presentation of an item) is precluded. However, once posited that $\prod p(a_i)$ is not a direct function of cognitive load, the model predicts variations of $\prod p(a_i)$ at constant cognitive load – $\prod odd(a_i)$ being constant, thus allowing a correct simulation of the cognitive load in the background, so $\prod p(a_i)$ could offer a better angle than the log-odds to account for the consolidation or refreshing effects, item per item. In other words, the TBRS2 can refine the predictions of the more stringent implementation of the TBRS mimicked by the log-odds.

For instance, suppose that the probability of recall of two items X and Y are .5 each. Then, the corresponding odd is 1. Therefore, $\prod odd(a_i) = 1$, whereas the probability of the combined correct recall is $\prod p(a_i) = 1/4$. On the other hand, suppose that $p(X) = 1/3$ and $p(Y) = 2/3$, then $\prod p(a_i) = 2/9$, but the odds, now being $1/2$ and 2 still multiply to 1. In fact, in the TBRS2 model, this translates to the observation that under constant cognitive load, the span will be higher when the activation of items are more similar. We can therefore explain small variations in probability of recall without the need to add a specific consolidation process. For instance, when an item cannot be refreshed right after its presentation, this might lead to a greater differential in the activations of the to-be-recalled items. Therefore, crucially, we believe that the so-called consolidation phenomenon can be explained by the mere effect of the refreshment process.

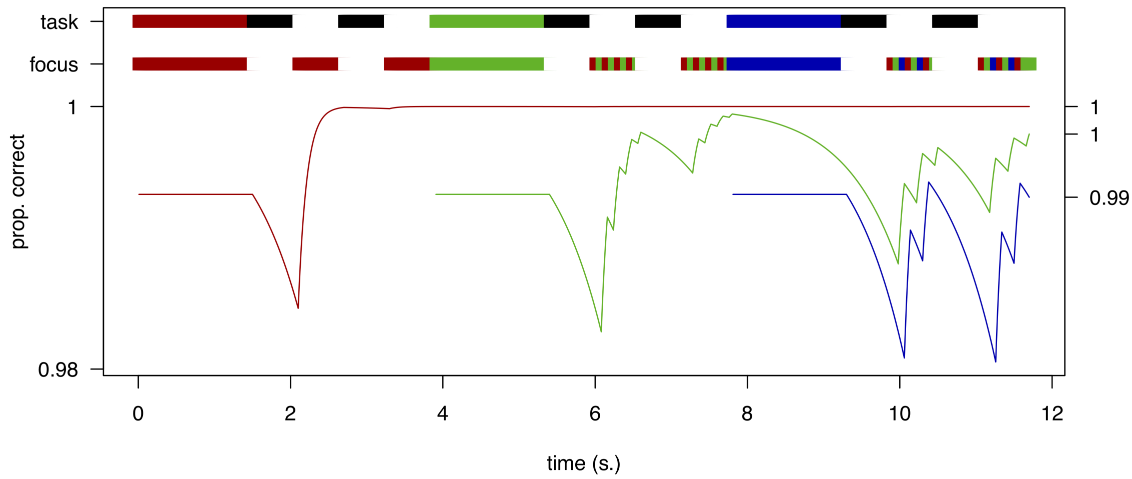
This explanation is not only more parsimonious (and therefore to be preferred, all other things being equal), but also can explain why the consolidation phenomenon is controversial, as the consolidation effect may appear in some circumstances, but not in others. The next section details numerically how this is possible in greater detail.

Let's consider as an example the two complex tasks "L1010L1010L1010" and "L0011L0011L0011" , where L stands for a letter presentation (1500 ms), 0 for a free period of 1000 ms, and 1 for a period of 1000 ms devoted to the concurrent task. The predicted probability of correct recall of the whole sequence of 3 letters is 99.10% in the first case, but 99.49% in the second case[2], although the cognitive load is 50% in both tasks. This difference could be interpreted as a consolidation effect in the second task (because the participants have a greater period of free time after each letter presentation). However, as can be seen in Figure 5.2, this is explainable as a consequence of the fact that the probabilities of item-wise correct answer are more scattered in the first case.

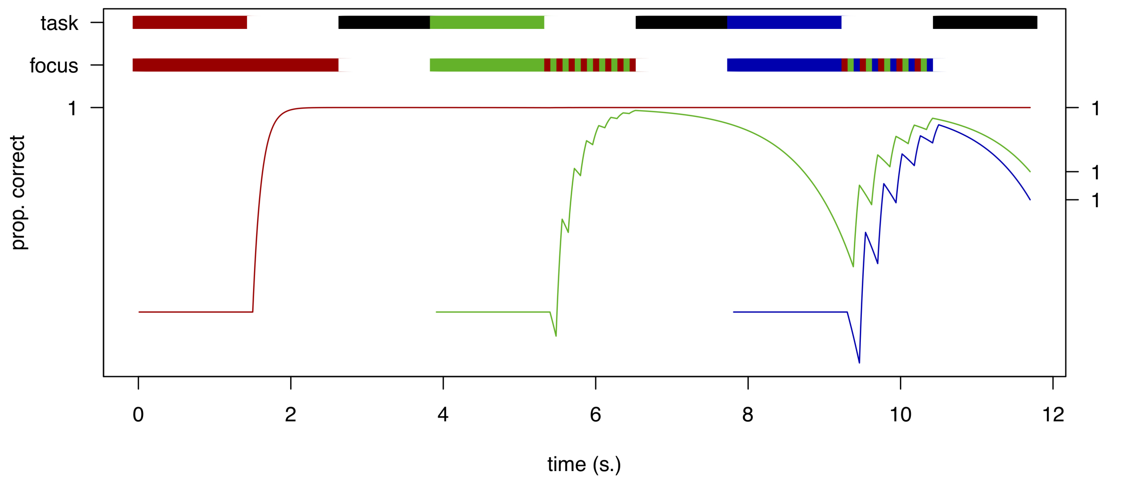
However, the product of odds are identical in the two situations (see Figure 5.3), amounting to 191,853,903,489. The item-wise odds are $2808047.6125, 475.3256$ and 143.7391 respectively in the first case, and $1222000.7103, 475.3256$ and 330.2996 in the second case. In both cases, the product of these three values amounts to 191,853,903,489. The numbers can appear huge for simulating a cognitive process, but again, more useful probabilities can be computed once the strict implementation of the model based on odds has produced its prediction. The trick is that the odds allow the model to be perfectly implemented while the probabilities can allow greater variations in the predictions of the model.

Put simply, the positional variation of the processing items can still produce different patterns of recall, even when the mathematical constraints of the model are thought to best fit the assumptions of the model. As explained above, this is the case

[2]This difference seems small, but can be increased when tuning the default parameters.

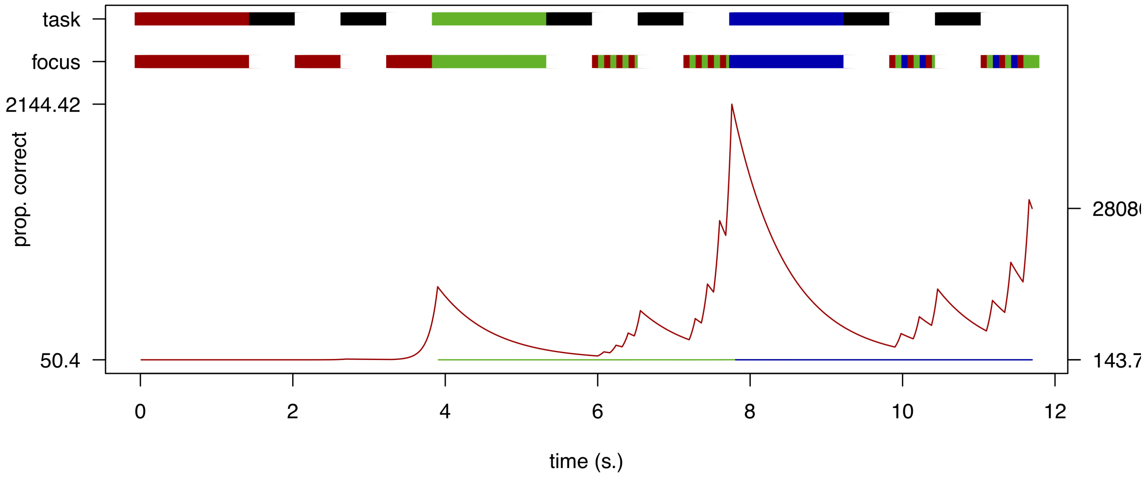


(a) Predictions for the task L1010L1010L1010.

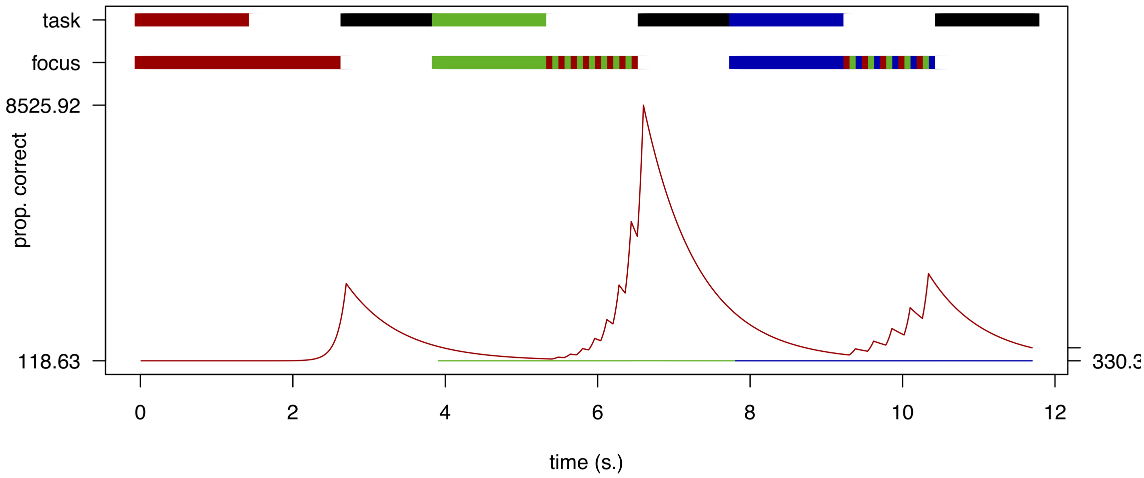


(b) Prediction for the task L0011L0011L0011.

Figure 5.2: The graphs displays the temporal evolution of the probability of recall of each item as predicted by the TBRS2 model. The end probability of overall correct recall is greater on the bottom panel, because the item-wise probability are less scattered. *Note*: the parameters were set as follow. Duration 80 ms, baseline 5, $d = 1.4$, $r = 9$, no switch



(a) Predictions for the task L1010L1010L1010.



(b) Prediction for the task L0011L0011L0011.

Figure 5.3: The graphs displays the temporal evolution of the odds of recall of each item as predicted by the TBRS2 model. *Note*: the parameters were set as follow. Duration 80 ms, baseline 5, $d = 1.4$, $r = 9$, no switch

in particular when long uninterrupted delays of free time are introduced in a sequence. However, the two types of tasks illustrated above (L1010L1010L1010 vs. L0011L0011L0011) are not ideal for testing the consolidation effect because the number of switches between the '0' and '1' events is doubled in the second task. Figure 5.4 describes two other different types of positional variations of the processing patterns to better test the consolidation effect.

In Figure 5.4, the letters L and K are examples of to-be-remembered items. The numbers in between the letters represent the processing events and the empty rectangles in the sequence correspond to free time. The only difference between the two different sequences A vs. B is whether the processing item is immediately presented after the storage item (thus creating a longer uninterrupted delay of free time later on in A) or whether there is a delay of free time immediately after the presentation of the letter (as in pattern B). According to the consolidation hypothesis, the task involving the free time after the letter should produce better performance in working memory compared to the task where free time is presented only after the first processing item (i.e. performance $B > A$). The rationale is that considering that consolidation is distinct from refreshing, consolidation should be most efficient immediately after the encoding of the storage item.

Interestingly, the TBRS2 as described above makes the opposite prediction. Increased working memory performance is expected when the processing event is presented immediately after the letter (i.e. pattern $A > B$) because, according to the TBRS2, a long uninterrupted refreshing opportunity is more beneficial than if this free delay is dispatched across the sequence.

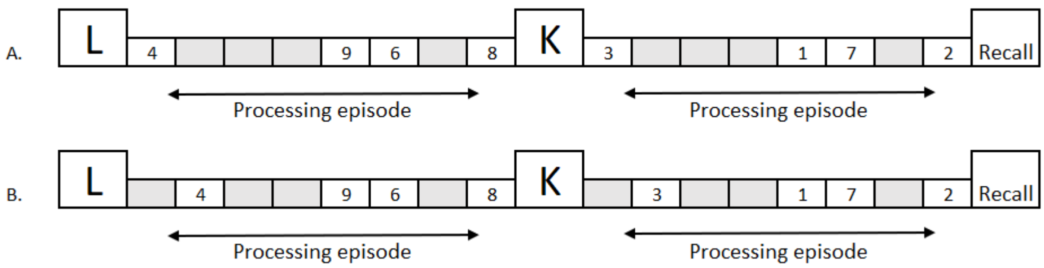


Figure 5.4: Patterns of the processing task in Experiment 7. *Note*. Pattern A provides no consolidation opportunity with a long uninterrupted opportunity to refresh after the first processing item. Pattern B provides an opportunity for consolidation, followed by a short opportunity to refresh after the first processing event.

Figure 5.5 shows the predictions of the TBRS2 for this case. As seen from the

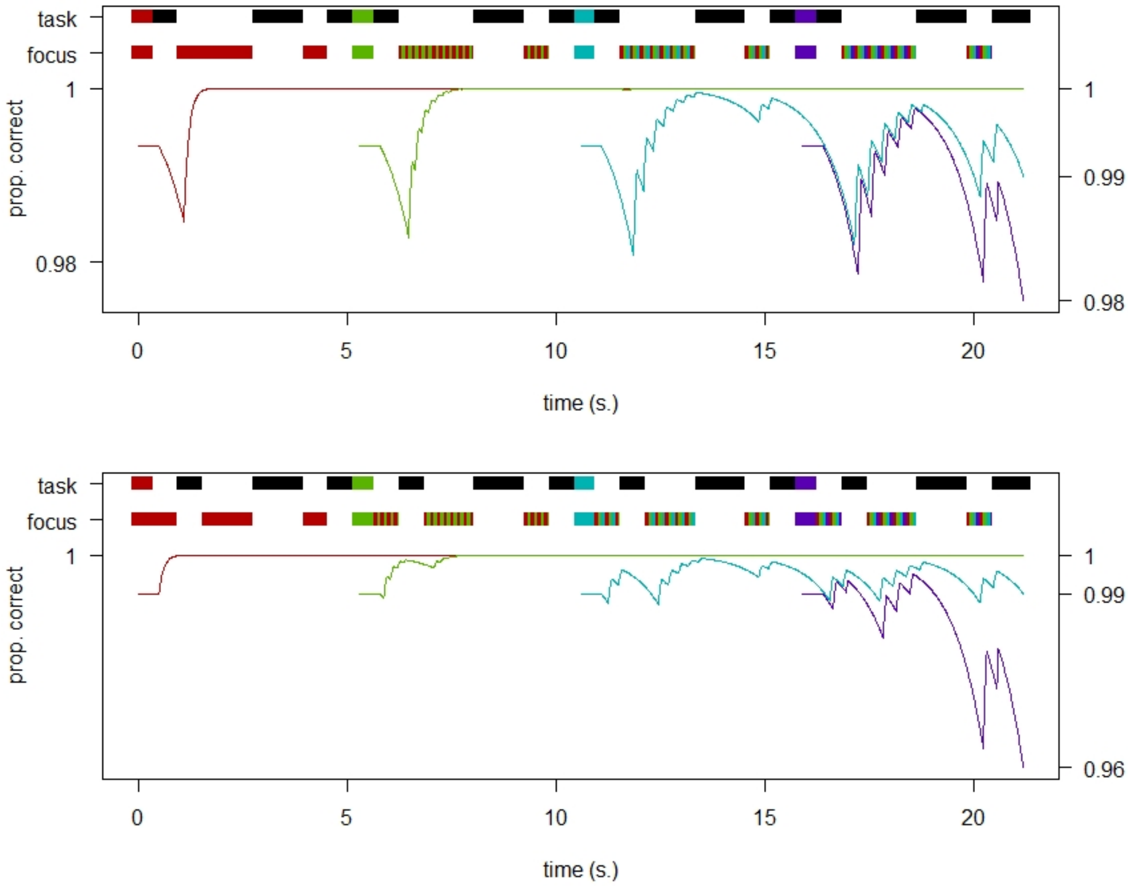


Figure 5.5: Top: Predictions of the TBRS2 for a sequence of four letters starting with immediately a processing episode : 1.0000000 0.9999719 0.9899293 0.9755888. Bottom:Predictions of the TBRS2 for a sequence of four letters starting with a consolidation opportunity. The probability of correct recall for the successive four letters are: 1.0000000 0.9999574 0.9933337 0.9634557 Both simulations involved a sequence of four letters and the parameters the parameters set as in Fig. 5.1

.

numbers in the figure's note, the probabilities of correct recall of the different items differ between the conditions and so is the product of these probabilities between the two tasks. When there is no opportunity to consolidate the memorandum (pattern A), the product of the probabilities of recall for the four items is 0.9657. When the opportunity to consolidate the memorandum is provided (pattern B), the product of the probabilities of recall for the four items is 0.9567. From the focus function (i.e. colored squares that display the storage and refreshing schedules), we can see that the refreshing schedules differ from one task structure to another. In the top figure the sequence ends with the turquoise color, which translates to a higher probability for this last item, whereas this is not the case for the last item in the bottom Figure. In contrast, in the bottom Figure it is the first three items that show higher probabilities of recall in comparison to the case in the top Figure. In essence, these item-wise probabilities of correct recall will overall predict different outcomes.

In the TBRS2, the gain can be explained by the amount of uninterrupted free delay, whether it is right after the storage item or not, instead of being directly linked to sustaining an item just encoded. Our hypothesis is, thus, that if performance increases in this case, then consolidation as a specific working memory maintenance mechanism is to be questioned, as refreshing and consolidation would appear to be less distinct than previously suggested. The alternative is the consolidation hypothesis, according to which free delay after the presentation of a storage event should lead to better working memory performance. In the next section, we attempt to verify our hypothesis and present the experimental results of these two different types of patterns of processing events in a complex span task.

5.2 Experiment 7

5.2.1 Method

Participants. Thirty-two students (27 females, 5 males; mean age = 22.76 years, $SD = 6.72$) from Université Côte d'Azur participated to this experiment in exchange for partial course credit. All participants were tested in a quiet room at the university. The duration of the experiment was approximately 40 minutes.

Material and procedure. Our within-subjects experiment manipulated the position of the processing events using two conditions: No consolidation opportunity (with a long uninterrupted free time for refreshing after the first processing item) *vs.* Consolidation opportunity (followed by a short window of free time to refresh after the first processing event). The two conditions were implemented respectively by the patterns A vs. B in Figure 5.4, which separated the to-be-recalled letters.

The patterns A and B both consisted of four digits (in their Arabic forms) randomly drawn from 1 to 9 and four slots of free time. After every event (i.e, letter, digit or free time) an interstimulus interval of 100 ms was presented. In detail, the pattern A consisted of a 100 ms interstimulus delay after the presentation of the storage item followed by the display of a digit for 600 ms and then a delay of 2200 ms. Then, two consecutive digits were each displayed for 600 ms and separated by a 100 ms interstimulus delay. Finally, a 800 ms delay was introduced before the presentation of the last digit for 600 ms + 100 ms interstimulus delay. The total amount of free time in pattern A once the storage item had been presented was $100 + 2200 + 100 + 800 = 3200$ ms. The pattern B involved a delay of 800 ms after the presentation of the storage item. Then, the first processing item was presented for 600 ms followed by 1500 ms of free time and two consecutive digits each displayed for 600 ms and separated by a 100 ms interstimulus delay. Finally, a 800 ms delay was again introduced before the presentation of the last digit for 600 ms + 100 ms interstimulus delay. The total amount of free time in pattern B once the storage item had been presented was $800 + 1500 + 100 + 800 = 3200$ ms.

Participants were invited to memorize in the correct serial order letters while reading aloud digits. The letters were randomly chosen from a permutation of the following consonants: B, F, H, J, K, L, N P, Q, R, S, T. Each trial started with a fixation cross displayed for 750 ms, followed by a delay of free time and a letter, each presented for 500 ms. After each consonant, either the processing pattern A or B was displayed. At the end of each trial, a matrix of all the possible consonants which could be presented during the experiment was displayed at the center of the computer screen. The participants were invited to recall the letters in their correct serial order by clicking on buttons representing the letters. Before starting a new trial, a feedback on the number of letters correctly recalled for the current trial was presented on the screen.

We used a memory load calibration task before starting the experimental trials. Given individual differences in working memory capacity (and previous result in Experiment 4 showing that consolidation was storage sensitive), this calibration process ensured that an optimal number of letters would be presented to the participants to avoid ceiling and floor effects. The calibration task started with three trials per processing pattern, with a set size of four letters interspersed with free delays and distractors. The order of the processing patterns in the calibration task was random without replacement. The processing schedules were the same as the ones presented in the experimental blocks. After completing six trials for a given set size, the set size was increased by one whenever the overall percentage of correct serial recall was 50% or above. Otherwise, the calibration task ended and the experimental trials started. This procedure was continued until the participant reached a correct serial recall rate inferior to 50%. Four training trials of a set size of two letters preceded the calibration task.

The experimental task was identical to the calibration trials, except that the individually calibrated set size remained constant across trials. Each pattern was presented in blocks of 15 trials, thus each participant completed a total of 30 trials. The order of the first pattern was counterbalanced between participants.

5.2.2 Results

We first analyzed the overall effect of the processing patterns on the proportion of correct recall in the correct serial position, referred to as memory performance from now on. Following our main analysis, we examined the relation between the type of pattern in the concurrent task and the serial position of the to-be-recalled letters. This second analysis was conducted with a Generalized linear mixed model. In this analysis, we included only the 23 participants (out of 32 in total) for whom the result of the calibration was four letters. To avoid an overestimation of the proportion of correct recall for the last items of the list by higher span participants, we split the data by set size for an analysis of serial position effects. Given that the majority of participants received a set size of 4, we focused only on these individuals. Both the main and the serial position analysis followed the same rationale as in Chapters 3 and 4. The statistical models and their ranking are presented in Table 5.1, Table 5.2, and Table 5.3.

Table 5.1: Statistical models of the main analysis created for Experiment 7.

$M0 \longleftarrow \text{DV}\sim + (1|Subject)$
$M1 \longleftarrow \text{DV}\sim \text{Processing Patterns} + (1|Subject)$

Table 5.2: Statistical models created for the serial position analysis in Experiment 7.

$M0 \longleftarrow$ DV$\sim$ + $(1|Subject)$
$M1 \longleftarrow$ DV$\sim$ Processing Patterns + $(1|Subject)$
$M2 \longleftarrow$ DV$\sim$ Serial Position + $(1|Subject)$
$M3 \longleftarrow$ DV$\sim$ Processing Patterns + Serial Position + $(1|Subject)$
$M4 \longleftarrow$ DV$\sim$ Processing Patterns * Serial Position + $(1|Subject)$

Table 5.3: Summary of model selection in Experiment 7 for memory performance and serial position performance based on the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. Summary of models in Table 5.1 and Table 5.2.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Memory performance				
$M0$	3	305.84	0.00	0.66
$M1$	4	307.18	1.33	0.34
Memory performance as a function of serial positions				
$M1$	5	2520.56	0.00	0.58
$M3$	6	2521.35	0.79	0.39
$M4$	9	2526.91	6.34	0.02
$M0$	2	2591.84	71.28	0.00
$M2$	3	2592.66	72.09	0.00

Table 5.4: Mean memory performance (SD) as a function of serial position of the letters to be recalled, split by the consolidation opportunity in Experiment 7.

$Consolidation$	1	2	3	4
No	.88 (.32)	.82 (.38)	.72 (.45)	.74 (.44)
Yes	.91 (.29)	.84 (.37)	.73 (.44)	.77 (.42)

The results of the main analysis regarding the global effect of processing patterns on memory performance showed that the null model was best ranked, suggesting no evidence in favor of the effect of the order of processing events on recall ($M0/M1 = 1.95$). The mean memory performance by processing patterns was 0.55 ($SD = 0.5$) and 0.56 ($SD = 0.5$) for the no-consolidation pattern (A) and the consolidation pattern (B) respectively. Table 5.4 provides the mean memory performance as a function of the serial position and the processing pattern. Our second analysis examining memory performance across serial positions confirmed this result. In this second analysis, the best model was the one that only included the serial position factor. The model including the processing pattern factor alone was ranked last. This finding suggested that memory performance was not enhanced by a consolidation opportunity, as seen in Figure 5.6.

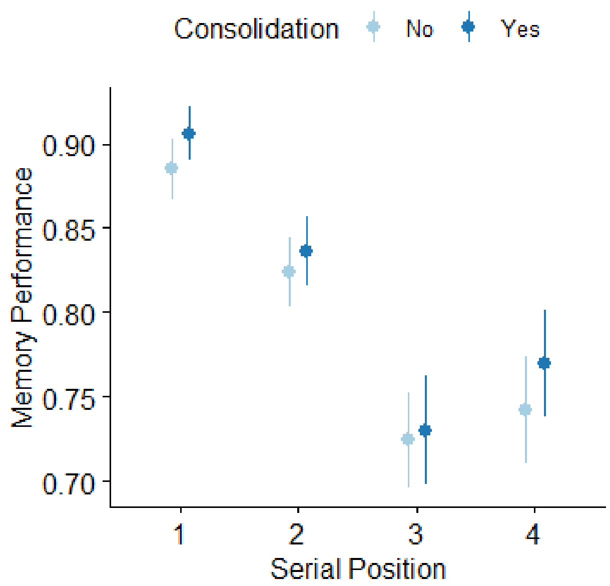


Figure 5.6: Mean memory performance as a function of serial position of the letters to be recalled, split by the consolidation opportunity in Experiment 7. *Note.* In the legend, No consolidation refers to pattern A in which a digit is immediately displayed after the storage event. The consolidation (yes) refers to pattern B in which a delay of free time is displayed before the presentation of the first processing item. Both patterns are represented in Figure 5.4.

In sum, the results suggest that the consolidation effect can be equated by introducing an uninterrupted long delay of free time after a first processing item which supposedly halts consolidation. However, the display duration of the free delay in the consolidation processing pattern (B) was only 600 ms, and it is possible that this duration was not long enough to allow consolidation to take place. Furthermore, because we did not collect concurrent task performance data, we were unable to examine if participants delayed their responses when no consolidation opportunity was provided, and in which case this could possibly explain why no difference was found between the two conditions. Experiment 8 addresses this issue.

5.3 Experiment 8

Experiment 8 was a follow-up to Experiment 7. The goal was to examine whether participants delayed their responses at the concurrent task because of incomplete consolidation. Additionnally, the duration of the consolidation opportunity was increased. The same patterns A and B in Figure 5.4 were used as in Exp. 7. We opted for a parity judgment task in order to collect RTs in the concurrent task.

5.3.1 Method

Participants. Twenty-seven students (27 females, 2 males; mean age = 21.93 years, $SD = 4.65$) from Université Côte d'Azur participated to this experiment in exchange for partial course credit. All participants were tested in a quiet room at the university. The duration of this experiment was approximately 1 hour.

Material and procedure. The material was identical to Experiment 7, but here the participants were required to complete a parity judgment task instead of a reading digit span task. Digits and interstimulus intervals were fixed at 900 ms, and the order of the events was determined by the same processing patterns as those used in our previous study. This pace was chosen to ensure that the consolidation opportunity in the Pattern B would be long enough for the post-encoding processes to take place. The procedure was identical to our previous study.

5.3.2 Results

In this Experiment, two participants were excluded from the following analysis as their performance at the concurrent task deviated more than two standard deviations from other participants. For the remaining twenty-five participants: mean age $= 22.09$ years, $SD = 4.72$.

Effect of consolidation on recall. We used the same scoring method of recall performance as in our previous experiments (i.e. proportion of correct recall in their correct serial position), referred to as memory performance from now on. We ran the same analysis as in Experiment 7. The main analysis examining the global effect of the processing patterns was conducted on all the participants. Akin to our previous experiment, the second analysis exploring serial positions as a function of the processing pattern was conducted only for the participants for whom the result of the calibration was seven letters ($N = 22$). The statistical models and their ranking are presented in Table 5.1, Table 5.2, and Table 5.5.

Overall, the results were similar to our previous experiment, showing no evidence in favor of consolidation. The mean memory performance by processing patterns was 0.76 ($SD = 0.31$) and 0.77 ($SD = 0.30$) for the no-consolidation pattern A and the consolidation pattern B [3]. Table 5.6 provides the mean memory performance

[3]The calibration task was limited to 7 letters, which explains why the mean memory performance

Table 5.5: Summary of model selection in Experiment 8 for trial-by-trial memory performance (top) and serial position accuracy (bottom). Model selection was based on the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. Summary of models in Table 5.1 and Table 5.2.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Memory performance				
$M0$	3	113.22	0.00	0.71
$M1$	4	115.05	1.83	0.29
Memory performance as a function of serial positions				
$M1$	8	4053.61	0.00	0.61
$M3$	9	4054.49	0.88	0.39
$M4$	15	4065.02	11.41	0.00
$M0$	2	4235.90	182.29	0.00
$M2$	3	4236.84	183.22	0.00

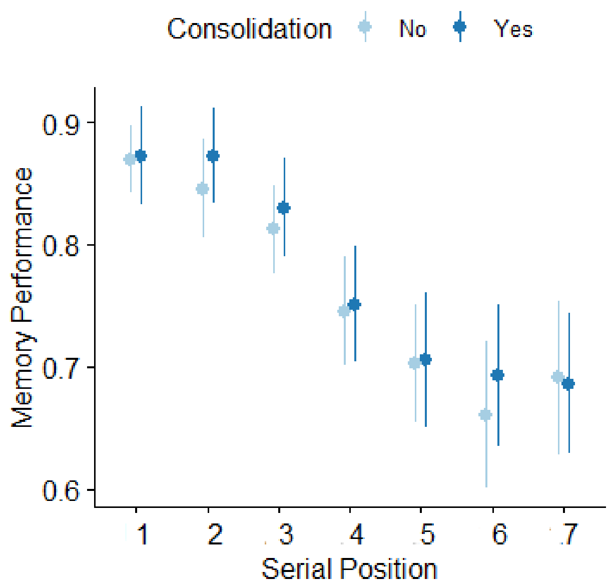


Figure 5.7: Mean memory performance as a function of serial position of the letters to be recalled, split by the processing patterns in Experiment 8. *Note:* In the legend, No consolidation refers to pattern A in which a digit is immediately displayed after the storage event. The consolidation (yes) refers to pattern B in which a delay of free time is displayed before the presentation of the first processing item. Both patterns are represented in Figure 5.4.

as a function of the serial position and the processing pattern. Both analyses ended up with a ranking of the models similar to our previous experiment. The null model ranked best for the analysis of the global effect of the processing patterns. Our second analysis examining memory performance across serial positions ranked the model including only the serial position factor as the best model. Again, the models that included the processing pattern factor were ranked last, which brings further evidence suggesting that the effect of consolidation opportunity on recall was equated in our experiments as seen in Figure 5.7.

Table 5.6: Mean memory performance (SD) as a function of serial position of the letters to be recalled, split by the consolidation opportunity in Experiment 8.

Consolidation	1	2	3	4	5	6	7
No	.87 (.34)	.85 (.36)	.81 (.39)	.75 (.44)	.70 (.46)	.66 (.47)	.69 (.46)
Yes	.87 (.33)	.87 (.33)	.83 (.38)	.75 (.43)	.71 (.46)	.69 (.46)	.69 (.46)

Effect of consolidation on concurrent task performance. We used the proportion of correct parity judgment as our dependent variable to examine the global effect of the processing patterns on the concurrent task. Similarly to our previous analysis, we also examined further the effect of the processing patterns on the serial position of the concurrent task (for participants that calibrated at 7 letters, $N = 22$). As our dependent variable, we used the mean proportion of correct parity judgment operations between two letters. See Table 5.2 for a summary of the statistical models, and Table 5.7 for their ranking obtained by model selection.

Our first analysis showed that the model including the processing patterns ($M1$) was best ranked with strong evidence in favor of the effect of the processing patterns on concurrent task performance ($M1/M0 = 105.05$). We observed a significant difference between the processing patterns on performance at the parity judgment task ($t(724) = -3.38$, $p < .001$). The estimate showed a decrease of 2% of the concurrent task performance when a distractor was immediately displayed after the storage item.

The results of the second analysis which examined the effect of serial position on the proportion of correct parity judgment operations confirmed this trend. We observed that the top-rank model was the full model without interaction ($M3$). Specifically, we observed strong evidence that both factors (serial position and processing patterns) contributed to performance at the concurrent task ($M1/M0 > 150$;

is higher than 50%

$M2/M0 > 150$). The model best ranked showed a significant effect of both factors on the parity judgment task ($F(6, 4591) = 22.89$, $p < .001$); $F(1, 4591) = 14.94$, $p < .001$ for serial positions and processing patterns respectively). The estimates of the mixed model indicated a respective increase of concurrent task performance as a function of the serial position of 1.5%, 0.7% and 0.2% for the positions 2, 3 and 4, in comparison to the first position; conversely, the estimates corresponded to a decrease of 4.4%, 7.4% and 9% for the remaining serial positions 5, 6 and 7. For the same dependent variable, the estimate for the processing pattern factor indicated a decrease of performance of 2.6% for the no-consolidation pattern (A) in comparison to the consolidation pattern (B). As seen from Figure 5.8a, the difference between the two conditions is mostly apparent on the first serial position, where performance was greatly reduced in the immediate condition compared to the delayed condition.

Table 5.7: Summary of model selection in Experiment 8 for performance at the concurrent task trial-by-trial, and performance measured as the proportion of correctly processed distractors position-by-position within sequences. Model selection was based on the number of parameters of each model (K), AICc values, delta AICc values and weight of each model. Summary of models in Table 5.1 and Table 5.2.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
Concurrent task performance				
$M1$	4	-1091.16	0.00	0.99
$M0$	3	-1081.85	9.31	0.01
Concurrent task performance as a function of serial position				
$M3$	10	-410.47	0.00	0.84
$M4$	16	-407.14	3.33	0.16
$M1$	9	-397.56	12.91	0.00
$M2$	4	-287.20	123.26	0.00
$M0$	3	-274.72	135.75	0.00

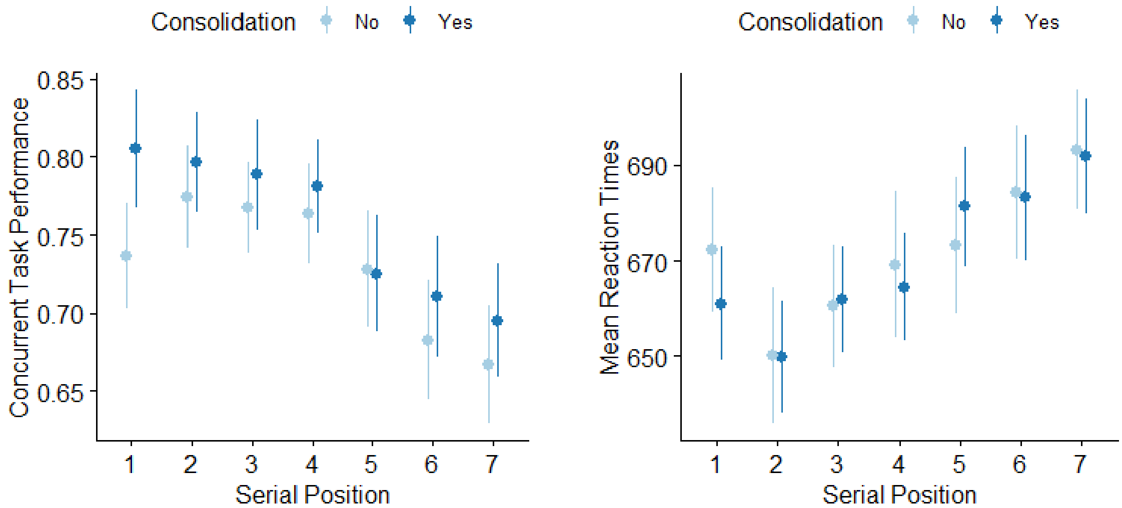
Effect of consolidation on RTs. We used the mean RTs for the parity judgment operations across serial position as another dependent variable to examine whether this measure would concur with accuracy on the concurrent task. We only included in the analysis RTs associated to a correct response.

Table 5.8 shows a ranking of the models. The results showed that the best model included only the serial positions as a significant predictor of the RTs ($M1$). The calculated evidence ratio between $M1$ and the null model suggested strong evidence in favor of the effect of serial position on RTs ($M1/M0 > 150$). In contrast, the model including only the processing pattern factor was ranked last, suggesting that

Table 5.8: Summary of model selection in Experiment 8 for serial position Rts of concurrent task based on the number of parameters of each model (K), AICc values, delta AICc values and weight of each model.

$Models$	K	$AICc$	ΔAICc	$AICcWt$
$M1$	9	2951.44	0.00	0.72
$M3$	10	2953.35	1.92	0.28
$M4$	16	2962.58	11.14	0.00
$M0$	3	3019.22	67.78	0.00
$M2$	4	3021.11	69.67	0.00

*Note: M0 = Null model; M1 = Serial Position; M2 = Processing Patterns; M3 = Serial Position + Processing pattern; M4 = Serial Position * Processing pattern.*



(a) Mean parity judgment task performance between two letters as a function of serial positions of the concurrent task, split by processing patterns ($N = 22$).

(b) Mean RTs for the parity judgments between two letters as a function of serial positions of the concurrent task, split by processing patterns ($N = 22$).

Figure 5.8: *Note:* In the legend, No consolidation refers to pattern A; Consolidation yes refers to pattern B. Both patterns are represented in Figure 5.4.

the presentation of a distractor immediately after the storage item did not influence RTs in the concurrent task. In detail, the results showed a significant effect of serial position on RTs ($F(6, 286) = 15.45$, $p < .001$). The estimate showed a decrease in RTs of 17 ms, 6 ms for the serial positions 2 and 3 respectively, in comparison to the RTs in the position 1, and an increase in RTs of 0.1 ms, 11 ms, 17 ms, and 26 ms for the 4, 5, 6 and 7 positions respectively. As seen from Figure 5.8b, RTs in the parity judgment task increased as the memory load of the task increased. In contrast, the RTs of the processing patterns did not differ much from each other across positions. Note that for both conditions, we see that RTs are longer for the first serial position than the rest of the serial position curve would expect. This could be interpreted as a restart cost.

5.3.3 Discussion

The current study examined whether the position of the first processing item in a complex span task could influence working memory performance. According to the consolidation hypothesis, a long delay before the processing task should produce better performance in working memory. The idea is that consolidation can be beneficial to recall because it supposedly stabilizes and strengthens novel information in working memory. An alternative hypothesis derived from the predictions of the TBRS2 model is that it is not the position of the delay per se that is beneficial but rather the amount of uninterrupted time enabling refreshing.

Both Experiment 7 and 8 confirmed that the consolidation effect on memory performance can be equated by introducing an uninterrupted long delay of free time after the first distractor. However, against the prediction of the TBRS2 model, the processing pattern with the long delayed opportunity to refresh did not produce better working memory performance compared to the pattern with an immediate opportunity to consolidate memoranda. Nonetheless, our results confirm the basic decay account of forgetting, as conceptualized by the original TBRS model. That is, equal memory performance can easily be predicted by the constant cognitive load across our experimental conditions (reflected by the proportion of attentional capture against free time devoted to refreshing). That is, the refinement of the computational model TBRS2 does not surpass the theoretical TBRS.

Regarding the results of Experiment 8 on the concurrent task performance, we observed that although the parity judgment task was executed worse when no consolidation opportunity was provided (i.e. pattern A), the RTs seem to suggest that the participants did not delay their responses. Nonetheless, it could be argued that worse performance at the parity judgment task could indicate that the consolidation process was interrupted thus causing difficulty to respond correctly to the first distractor. However, the absence of task preparation in the no consolidation processing pattern (i.e. B) could explain these results just as well.

One limitation of our study is potentially the fact that unlike in some experiments testing the consolidation effect (e.g. Bayliss et al., 2015; De Schrijver & Barrouillet, 2017), we did no use a mask. However, as established by Nieuwenstein and Wyble (2014), even though a mask can hinder the beneficial effect of consolidation, it does not stop it. Consolidation is interrupted only when attentional processes are focused towards the secondary task. The fact that we did not use a mask should have resulted in a beneficial effect in working memory performance as no additional interference was introduced. This was not the case. Another limitation is the lack of power. Given the subtle variations in the probability of correct recall predicted by the TBRS2 between the processing patterns with and without consolidation, the results of both experiments are arguably too subtle to provide a meaningful test of our hypotheses. While theoretically-relevant null results provide important information, there are still many reasons why an experiment produces a null outcome. In line with this reasoning, a first power analysis revealed that 750 participants would be necessary to achieve a power of .80 to detect a difference between our two main conditions. To verify whether a larger sample could potentially increase power to detect an interaction between the consolidation manipulation and position of the items, we run a power analysis for generalised linear mixed models using the simR package in R. The simulation based on our mixed model revealed that we should at least run 350 participants to achieve a power level of .80. However, these values exceed our pool of participants who could obtain course credits in exchange of their participation the year we finished conducting the study in 2020.

Overall, as we established in our introduction, recent studies on consolidation have been numerous. However to our knowledge, only two published papers have argued in favor of an effect of consolidation within complex span task paradigms (Bayliss et al., 2015; De Schrijver & Barrouillet, 2017), and only one argued that consolidation distinguishes from other maintenance processes (Bayliss et al., 2015). We discussed both of these studies in our introduction. The present studies attempted to provide an alternative account for the consolidation effect. Although we were not successful in providing evidence in favor of the TBRS2, we still believe that the model yields interesting alternative to explain some of the discrepancies found in the present dissertation and in the literature.

5.4 Conclusion

The aim of this Chapter was to test an alternative explanation of the consolidation effect derived from the predictions of the TBRS2 model. According to the consolidation hypothesis, the presentation of a free delay after a storage should benefit working memory performance. Supposedly, this free delay should be beneficial to memory performance because it stabilizes and strengthens novel information in working memory. An alternative hypothesis derived from the predictions of the TBRS2 suggests that when an item cannot be refreshed right after its presentation, the situation leads to a greater differential in the probability of correct recall of the to-be-recalled items.

We conducted two experiments in order to test these alternative hypotheses. In short, our results showed null results in both experiments, indicating equal memory performance with or without consolidation opportunity. Given our small sample size, these results are arguably too inconclusive to provide a meaningful test of our hypothesis. However, when null results are consistently found across a suitable range of plausible scenarios, they can be convincing. As such, these experiments provide information to better distinguish the ideas of consolidation and refreshing, with the aim of conceptualizing more minimalist models of working memory, but in definitive further studies are needed before the predictions of the TBRS2 model can be validated.

6 General Discussion

The present dissertation examined the psychological effects of structural variations of the complex span task to explore predictions of the Time-Based Resource-Sharing model (TBRS). For this purpose, we manipulated the global structure of a complex span task by varying the order of storage and processing cycles, or alternatively we manipulated only the processing structure of the task by varying the order of distractors and free delays in between them. The main motivation for conducting these experimental manipulations was to detect switching costs associated to working memory performance. We organize the following discussion around this main question of interest. In the first section, we will review the central premises of our work. Here, we will examine our conception of switches associated to the storage and processing activities viewed as respective task sets. Then, in the second part of this discussion we will attempt to draw a comprehensive account of our findings. This will be achieved by reflecting on the potential role of several maintenance mechanisms known in the literature, namely elaboration, rehearsal, and refreshing. Finally, we will conclude this work by evaluating the limits and perspectives.

6.1 Reexamination of the premises

Our main question of interest focusing on detecting switch costs in working memory was examined in five experiments by manipulating the structure of the processing schedules of a complex span task. Despite these several attempts, the body of evidence in favor of switch costs between storage and processing remains scarce. In detail, only one experiment showed a detrimental effect of switching on recall performance (Exp. 6b) and on processing speed in the concurrent task (Exp. 6a). In line with our initial hypothesis, we interpreted this finding as evidence in favor of executive processes occurring between storage and processing activities (i.e. the two main functions of working memory). Despite using similar experimental designs, the remaining studies failed to find such costs associated to complex span task performance (Exp. 2-5). So far, the lack of strong evidence in favor of our hypothesis has been discussed within the different chapters of the present dissertation essentially in terms of the potential role of methodological variations and confounds. However, in light of these general results, the premises of our hypothesis might need to be questioned. In this regard, our conception of switches associated to the storage and processing activities viewed as respective task sets will be be examined in the following sections.

6.1.1 Task set

The first premise of this work was to assume that the storage and processing activity of a complex span task is associated to a specific task set. The tasks typically used in the task-switching literature involve stimuli and responses that afford both tasks (e.g. a parity and magnitude judgment task requiring the same keyboard responses for a common set of digits). In contrast, in a complex span task, the storage and processing activity do not share the same stimuli nor the same response set. In fact, the storage activity do not require any *overt* responses from the participant until the end of a trial, when there is no longer a temporal overlap between the two types of tasks. Therefore, it could be questioned whether this function can really be considered as a task that is associated to a specific task set.

As noted by several authors, the definition of what precisely comprises a task in itself is difficult to establish (e.g. Kiesel et al., 2010; Monsell, 2003), but a broad definition suggested by Kiesel et al. (2010, p. 850) is that "(...) tasks entail performing some specified mental operation or action in response to stimulus input." This means that the storage activity that do not exhibit overt behavior can be conceived as a task as long as participants aim to achieve a specific goal state. In theory, the mere intention to do a task should activate the corresponding mental task set to achieve the goal in question (Kiesel et al., 2010). The distinction between tasks and task sets has been studied in situations where nominally different tasks might involve the same task set. For instance, Schneider and Logan (2005) showed using computational modeling that the task-relevant knowledge of a parity and magnitude judgment task can be retrieved from memory from a common task set. The opposite scenario has also been discussed in the literature where one task can involve different task sets. For example, according to Schneider and Logan (2014) a parity judgment task can be achieved either by retrieving information from memory (i.e. performance based on the knowledge that $2, 4, 6, 8$ are even digits and $1, 3, 5, 7, 9$ are odd digits) or by using arithmetic skills (i.e. the participant check if a digit is even by dividing it by 2). A fairly common assumption has been, however, to posit the existence of two different task sets for two different tasks.

What is important for our argumentation here is that the long-standing debate in the task-switching literature has not been to question altogether that the execution of a task requires a task set, but rather the continuous debate on this issue has focused on whether different parameters (e.g. stimulus-category rules, orientation of attention, response-stimulus mappings, response modality) should or should not be included within the concept of the task set (Vandierendonck, Christiaens, & Liefooghe, 2008; Vandierendonck et al., 2010). It goes without saying that not all tasks call upon every single one of these parameters when they are executed (Vandierendonck et al., 2010). Further arguments in favor of the idea that the storage activity is indeed mapped to a specific task set is provided by Monsell et al. (1996) who clearly states that single-step tasks (i.e. such as object naming or word reading involving a chain of processes that are mapping input to output; opposed to multi-step tasks such as making coffee) are associated to various control processes including task-set reconfiguration. Additionally, the author specifies that: "(...) single-step tasks also include the harder-to-study cases where the 'input' is internal: i.e. an idea, a memory, or the internal production of some other operation, evokes an act. And there are tasks where an input evokes no overt response–as when our response to an experience is merely to 'read, mark, learn, and inwardly digest' "(pp.95-96). The question is thus not whether the storage activity is associated to a specific task set, since this point is hardly disputable, but rather whether the competition (in terms of the reconfiguration or interference account) between the storage and processing task sets guiding complex span task performance is enough to produce switch costs.

Admittedly, there is little overlap and conflict between the two task sets guiding complex span task performance, since neither the stimuli nor the responses are shared between the two tasks. Although, one experiment provided some evidence in favor of the existence of switch costs between the storage and processing activity, most of our results did not reveal such costs. Previous task-switching experiments using univalent stimuli[1] have found compatible results with ours. These studies often report smaller switch costs when task switching occurs between univalent stimuli compared to bivalent stimuli or even cost-free task switching with univalent stimuli after extensive practice (Allport et al., 1994; Jersild, 1927; Rogers & Monsell, 1995; Ruthruff, Remington, & Johnston, 2001; Spector & Biederman, 1976; but

[1]A univalent stimuli afford only one task. In other words, the stimuli is specific to one task. For instance, in the complex span task the stimuli are univalent because letters are only associated to the storage activity, whereas the digits are associated only to the processing task. In contrast, a task with bivalent stimuli has the same stimuli in two different tasks (e.g. digits for a parity and magnitude judgment task).

see Strobach, Liepelt, Schubert, & Kiesel, 2012 for persisting switch cost for the RT data – but not for the error data – using univalent stimuli and responses despite extensive practice). Although we used in our experiments a subsequent number of trials of the complex span task, which could be considered as extensive practice, this alone cannot explain our results. Switch costs were also absent when looking at shorter list lengths that were executed in the beginning of the experiments where participants should have less practice. This leaves the limited competition between the two task sets as a more likely explanation for our results. This could mean that the presumed switch costs between storage and processing activities in our experiments are likely small to begin with, and that it could be neutralized when specific maintenance mechanisms, such as rehearsal, are deployed. (The role of rehearsal in our findings will be discussed below in section 6.2.2.)

According to Vandierendonck et al. (2010), the reduced switch cost found with stimuli that are specific to only one task is explained by both the reconfiguration and the interference account. The two task sets are simply so different from one another that they can hardly influence each other. Therefore, performance could be only stimulus-based with two active task sets guiding performance at the same time. In contrast, with overlapping task sets, time is needed either to reconfigure the upcoming task or to resolve the interference caused by the activation of the previous task set. Both accounts are thus compatible with our results, according to which the task sets guiding complex span task performance are perhaps too different from one another to produce substantial switch costs.

The above arguments are in line with the theoretical framework that has guided this dissertation. However, a parallel line of research has also investigated somewhat similar switch costs than those observed in the task-switching literature, as argued by some authors (Risse & Oberauer, 2010; Verschooren, Schindler, De Raedt, & Pourtois, 2019). In this parallel line of work, switch costs have been explored sometimes without making any explicit references about the construct of task set and its role in the causation of switch costs. The switch costs result in this case not from *task* switching but rather from *mental* shifting[2] (e.g. Garavan, 1998; Hedge & Leonards, 2013; Janczyk, Wienrich, & Kunde, 2008; Oberauer, 2002), although this distinction in the terminology is not always explicitly made (Ravizza & Carter, 2008). For instance, the study of Garavan (1998) exploring mental shifting

[2]Also called attentional, set or internal shifting/switching. The preferred terminology varies from one author to another, and overall, the ability to switch from one goal state to another (whether it involves nominally one or several tasks) has been referred to by various terms that are sometimes used interchangeably.

demonstrated the existence of object-switch costs between two running counters. In two experiments using a running count task, participants were required to count separately the number of appearances of two objects (i.e. squares and triangles) that are sequentially presented on the screen. The results showed that switching and updating a counter is associated to a cost (i.e. longer RTs) compared to just updating the same object. This finding was interpreted as evidence in favor of different states of accessibility of the representations in working memory. In other words, the cost is thought to reflect the time needed to shift attention within the memory set from one object to another.

Since there is no task change in mental-shifting experiments, such as in Garavan's study, it could be argued that in this specific context, switch costs do not result from the reconfiguration of the ongoing task set nor from the resolution of the interference of the previous task set. From this, we could perhaps extrapolate that the assumption of competing task sets is not a prerequisite for the detection of switch costs in complex span task performance. Simply, the assumption of a specific switching mechanism that controls the limited focus of attention as stated by Garavan (1998) could in theory suffice to potentially detect switch costs. Then, the act of shifting attention from processing to the storage object or the other way around could yield such costs. Of course, the fact that the object-switch costs described in Garavan's study are observed between two internal sets (as opposed to an externally or physically present stimuli) could question whether we talk about a fundamentally different type of switching than what is thought to occur in standard task-switching experiments, where task and response rules are presumably triggered mainly by an external stimulus (Verschooren, Schindler, et al., 2019). Of particular relevance to the present purpose, one could argue just as well that the switches thought to occur in a complex span task involve somewhat different mechanisms than what is observed in the standard task-switching or mental-shifting experiments, since the competition between the two main functions of working memory arises from both the external and the internal environment (i.e. stimuli that are physically present and that need to be processed *vs.* maintained representations in working memory, respectively). In sum, neither framework fits perfectly the switches that are thought to occur in a complex span task.

Recently, Verschooren, Schindler, et al. (2019) brought together the standard task-switching and the mental-shifting frameworks by asking whether we can expect to find switch costs when attention shifts from working memory representations to external information processing. To this end, the authors designed a sophisticated novel paradigm in which participants executed a visual discrimination task, considered as the baseline task. Occasionally on some trials, participants had to switch their attention to another external or internal task (i.e. visual search task *vs.* task involving the comparison of targets with a previously memorized set). Importantly, the material and the demands of these deviant trials were similar whether the task involved external or internal attention, allowing to compare the switches from different environments to the baseline (i.e. deviant external task → baseline task *vs.* deviant internal task → baseline task). The results showed that attentional shifts from internal to external information processing produced switch costs of similar magnitude than when these shifts occurred within the external environment. According to the authors, the observed switch costs reflect a bottleneck in attention control regardless of the source of information processing, which is compatible with the idea of resource sharing. As discussed by the authors, the results fit also previous work conducted within the neuroscience literature related to the gateway hypothesis (P. W. Burgess, Dumontheil, & Gilbert, 2007; P. W. Burgess, Gilbert, & Dumontheil, 2007). This hypothesis suggests that a control system based in the rostral prefrontal cortex might function as a gateway between external and internal information processing. Importantly, this gateway control system is not thought as the region of information processing, but as the routing system that balances the flow of information coming from external and internal sources (Henseler, Krüger, Dechent, & Gruber, 2011).

Following this study, Verschooren and his colleagues conducted two additional studies that used probe-to-target matching tasks to explore the existence of switch costs occurring when attention shifts from external to internal information processing (Verschooren, Liefooghe, Brass, & Pourtois, 2019; Verschooren, Pourtois, & Egner, 2020), as this was not explored in the aforementioned experiments. In these probe-to-target matching tasks, participants were required to compare a figure (i.e. probe) to a target which was either perceived or retrieved from memory. Again, the authors found reliable switch costs between internal and external information processing. Additionally, the switch costs appeared to be asymmetrical. Larger costs were associated to switches occurring from the external to the internal environment compared to the other way around, respectively estimated at 49 ms and 9 ms (Verschooren, Liefooghe, et al., 2019). According to the authors, this asymmetry was best explained in terms of associative interference occurring between previously learned stimulus features and competing attentional states.

Overall, the studies described here provide evidence in favor of switch costs occurring from shifts of attention between external and internal information processing. Nevertheless, the generalization of this effect to complex span task performance cannot be taken for granted. There is indeed many differences between the tasks involving internal attention in the studies by Verschooren and the storage activity of a complex span task. In the previously described experiments, the authors mention the following components as being probably involved in the execution of the internal task: stimulus detection and identification, retrieval of the memory load from working memory, comparison between the perceived stimuli on the screen and the retrieved representation, response decision and its execution (Verschooren, Schindler, et al., 2019). In the storage activity of a complex span task, we can hypothesize that the same components, with the exception of the comparison part, could occur. Based on these differences alone we can speculate several explanations for our lack of indisputable results. First, perhaps the requirement of recognition memory in the probe-to-target tasks (in contrast to the immediate serial recall in our complex span tasks), could imply different modes of working memory, at least to some extent. Previous studies have suggested that recognition and immediate serial recall rely on different encoding and retrieval processes. For instance, recognition is thought to benefit from encoding processes that make stimuli more different from one another, whereas immediate serial recall is enhanced by processes that boost serial order such as associative and organizational processes (Bhatarah, Ward, Smith, & Hayes, 2009; Derks, 1974; Hunt & Einstein, 1981; Tversky, 1973). These differences in the encoding and retrieval processes could potentially affect the amount of time needed to either resolve interference from the previous task, or the time needed for reconfiguration processes to take place. This could in turn affect the magnitude of the observed switch costs, as discussed previously.

Another, crucial difference between the storage activity and the internal task of Verschooren and colleagues, is that the implementation of the aforementioned processes are dispatched across a trial in a complex span task, whereas they are executed in an uninterrupted chain of processes in the internal task. Therefore, we could wonder whether or not the potential costs associated to switching are actually constant within a trial of a complex span task. If not, this could potentially mean that even though we presumably manipulated the number of switches from one trial to another, not all switches have an equally detrimental effect on performance. Somewhat related to this idea, is the finding that RTs decrease significantly after the first processing steps (Thalmann, Souza, & Oberauer, 2019; Vergauwe et al., 2014). This finding has typically been interpreted as evidence that the longer RT of the first processing item reflects consolidation processes, whereas the subsequent ones reflect maintenance processes (Bayliss et al., 2015; De Schrijver & Barrouillet, 2017; Vergauwe et al., 2014). However, for other authors this result could be attributed to a switching cost in working memory, because longer processing times for the first processing item are also found in the absence of a memory load (Oberauer & Lewandowsky, 2011). In line with this latter interpretation, it could be that orienting attention from one object to another becomes more automatic with the practice of several consecutive processing steps within a single trial. An alternative explanation combining the switch cost and the consolidation account could consider additive effects of these two processes. In this case, increased costs are produced in the initial stages of encoding when specific processes are thought to occur (e.g. consolidation). In contrast, they could be almost nonexistent when the maintenance activity takes place in the absence of a physical stimuli. In opposition to the dispatched processes occurring in a complex span task, in the studies by Verschooren, the observed internal switch costs reflect many processes that occur in a burst without interruptions. Additionally, the switches between the internal and external information processing activities occur well after the initial encoding processes have already taken place. In that regard, it could be argued that the internal task reflects long-term memory and not working memory activity, although the frequent requirements to retrieve the memorized information speaks in favor of the involvement of working memory (Verschooren, Schindler, et al., 2019).

Finally, other differences between the storage activity and the internal task could be related to how switch costs are measured. It could be that the probe-to-target tasks allow a more “online” or immediate measure of the potential costs. In contrast, the many interruptions that are characteristic to the complex span task might not allow to capture such costs because they could be too transient in nature, at least for the memory task. This could mean that despite the putative switch cost participants could manage to reach back their span level because they have enough available time to do so. Closely related to this point, the dependent variable is likely to influence the detection of switch costs. Collecting RT data is in all likelihood a more sensitive measure than the correct recall data that we collected in all of our experiments. In line with this idea, the study by Strobach et al. (2012) showed persisting switch cost after extensive practice for the RT data but not for the error data. In our work, we did collect and analyze RT data for the concurrent task in Experiments 4, 6a and 6b. The results indicated a detrimental effect of switching for the RTs only in Exp.6a. However, participants were also more accurate in this case, which could imply that participants made strategic speed-adjustments. This leads to question whether assessing latency scores alone for the concurrent task as we have done in Exp. 4 and 6b. is a valid measure of switch costs. The fact that we found beneficial effects of switching in these two experiments bolsters our suspicions that the latency scores alone are perhaps far from ideal because we are unable to assess whether strategic adjustments were also made by the participants in these studies. At the same time, some authors have also argued that assessing separately latency and accuracy scores for the detection of switch costs, as we have done in Exp. 6a, is also problematic because it does not allow to reveal the proper effect of switching on task performance. For instance, Hughes, Linck, Bowles, Koeth, and Bunting (2014,p. 705) argued that: "Separate analyses essentially ignore one reflection of cost (either latency or accuracy) at a time. If some individuals tend to exhibit costs in one variable more than in the other or even switch between experiencing more cost to accuracy or to speed of performance, the effect that the researcher is attempting to detect will be weakened and, therefore, more difficult to detect, when speed and accuracy are analyzed separately." To remedy this issue Hughes et al. (2014) recommended different scoring methods that combine latency and accuracy into a single reliable score for task-switching studies. One of these methods is a rank-ordering binning procedure in which performance on switch trials is compared to performance on non-switch trials. Although the alternative scoring methods discussed by Hughes et al. (2014) are certainly more reliable than using only latency or accuracy scores alone, our experimental design did not incorporate a pure block of non-switch trials (i.e. concurrent task without memory load). Therefore, the reanalysis of our data with these methods are unfortunately not possible. Adding such a block in our future experimental designs examining switch costs in complex span tasks could be easily achieved, and this could perhaps help overcoming some of the limitations we have described above.

In summary, this section reviewed the first premise of our work which concerns the idea that storage and processing activities are associated to specific task sets. We established that according to the task-switching literature, the absence of switch costs between the two main functions of working memory could be explained by the lack of competition between the task sets guiding complex span task performance. On the other hand, we also discussed the existing parallels between switching in a complex span task and attentional shifting between external and internal information processing. Recent work revealing switch costs between different sources of processed information proved to be of particular relevance to our own hypothesis. Numerous caveats prevented us, however, from drawing definitive conclusions on the existence of similar costs occurring in complex span tasks.

Nevertheless, this recent prior work along with our own –limited– results provide additional arguments in favor of the idea that switch costs between storage and processing could be the logical consequence of a single content-general attentional resource. Subsequently, the detection of switch costs in complex span tasks could perhaps boil down to a question of measurability. With this in mind, we turn to the second premise of this work which examines this issue of measurability through our conception of the switches between the two main functions of working memory.

6.1.2 Switches

The second premise of our work relates to our conceptualization of the switches. According to the TBRS model, working memory functioning involves a constant alternation between storage and processing activities through a rapid switching mechanism. The role of switching is to avoid the complete loss of information by allowing to refresh information whenever possible. Hence, switching is intimately related to the maintenance of the memory traces within the TBRS model. In our introduction, we hypothesized how these switches between storage and processing activities might occur during a complex span task. We suggested two different accounts derived from the TBRS(2) models. The main difference between these accounts boils down to when refreshing occurs. In the following section, we will reexamine these conceptions and assess the evidence in favor of each view as well as discuss the theoretical implications. To ease comprehension of this next section, we refer the reader to Figure 1.17 and Figure 1.18 of pp. 69 and 71 of Chapter 1 for a visual aid of the two accounts of switching between storage and processing activities.

TBRS vs. TBRS2 conception of the switches between storage and processing.

Throughout this work we endorsed our TBRS2 conceptualization of the switches (represented in Figure 1.17a). According to this view, the switches between storage and processing are governed by the task structure. Essentially, this view translates to count a switch each time a distractor follows a free delay or the other way around. The TBRS2's view of the switches can be derived to consider that the cognitive load function already accounts partially for the switch costs in working memory (represented Figure 1.17b). Because the cognitive load of a task comprises the proportion of time during which attention is not directed towards the memoranda, this leaves only the switches from processing to storage unaccounted for. This second view of the TBRS2 translates to count a switch only when a free delay follows a distractor (and not the other way around). Although the mechanics of counting the switches are slightly different, the main idea of the TBRS2 is that the switches are governed by the structure of the task. Therefore, the behavioral predictions of these two views are the same. For both views, we expected impaired memory and concurrent task performance as a function of the number of switches occurring in a trial. The original model's view differs from ours. The main difference between the two conceptualizations is related to when refreshing is thought to occur. The original model assumes continuous switching between storage and processing. This continuous switching process is also paired with a high-speed refreshing mechanism which allows to reactivate the degraded memory traces. Subsequently, switching can occur before and after *each* processing step, but also during the concurrent activity by taking short pauses in order to refresh the memoranda (Barrouillet & Camos, 2014b; Camos, 2017). Essentially, the original TBRS model's view means that every processing step is coupled with a switch from processing to storage and the other way around (see Figure 1.18a). Subsequently, the number of switches from storage to processing (and the other way around) is always equal, no matter the structure of the task. This account predicted that our manipulations of the switches should not influence working memory performance. It is worth noting that even if we consider that cognitive load already captures the switch costs arising from storage to processing, the predictions of the original TBRS model regarding our experimental manipulation remained unchanged.

Evidence & theoretical implications for the switching mechanism.

In favor of the original TRBS' predictions (and against ours), four experiments revealed no detrimental effect of our manipulation of the switches on working memory performance (Exp. 2-4). In detail, both Exp. 2 and 3 showed no effect whatsoever of our manipulation on recall or concurrent task performance. In experiment 4, the results showed that the structure of the processing task can incur slight variations of performance in working memory. However, when we attempted to isolate the patterns that manipulated best switching according to our view (i.e. dropping out from the analysis patterns that tapped into consolidation or irregularity), we found that the effect of switching was not statistically significant for recall performance. In contrast, the results of the concurrent task for this same experiment showed that performance was actually enhanced by our manipulation of the switches (i.e. faster RTs as a function of the number of switches occurring in a trial). The TBRS(2) models do not predict such an effect of switching. Therefore, we suggested that this unexpected result could reflect other mechanisms prompted by the structure of the task (e.g. task preparation and task predictability). Another interpretation is that the beneficial effect of our manipulation of the switches reflects speed-accuracy adjustments made by the participants (as discussed in section 6.1.1).

Taken together, the results of Exp. 2-4 show that, at least for recall, our manipulation of switching between storage and processing (or the other way around) do not influence performance. Hence, these findings suggest, in line with the original model, that switch costs between the two main functions of working memory can only be tacitly assumed at best, but it is not possible to quantify them by varying the structure of the task. However, the results of our last experiment cast doubt on this interpretation, as discussed below.

In favor of our TBRS2 predictions, one experiment provided some supporting evidence for our hypothesis, according to which our manipulation of the switches between storage and processing incurs a detrimental effect on working memory performance. In detail, Experiment 6a showed slower (but more accurate) concurrent task performance as a function of our manipulation of the switches between storage and processing (and the other way around). However, this same experiment showed no effect of our manipulation on memory performance. In contrast, credible evidence in favor of a detrimental effect on recall stemming from our manipulation of the switches was found in Experiment 6b. In this experiment, our manipulation of the switches was efficient in particular for the condition involving a low cognitive load (induced by an easy concurrent task) and a slow pace. Regarding the concurrent task data of this same experiment, we again found that performance was enhanced by our manipulation of the switches (i.e. faster RTs).

The most straightforward implication of our positive result is that variations of the structure of the complex span task can induce slight variations of working memory performance. Furthermore, the finding of impaired memory performance as a function of our manipulation of the switches in Exp. 6b indicates that switch costs can occur (at least in some cases) independently of the attentional capture generated by the processing task. Consequently, this finding brings some support to our TBRS2 conceptualization of the switches, suggesting that participants did not switch to maintenance activities when several processing steps are presented in a row (in Exp. 6b). However, our results do not allow to determine whether the observed detrimental effect is caused only by switching from processing to storage, or whether the switches occurring from the opposite direction also contribute somehow to the finding. As a consequence, the question of whether all switches are equal remains open-ended.

Overall, our results favored the original TBRS model's predictions according to which switch costs between the two main functions of working memory cannot be quantified by manipulating the structure of the task. One question remains however unanswered, why did we find evidence in favor of a detrimental effect of our manipulation of the switches on memory performance in one experiment? Should we consider this result simply as a false positive? If we examine the balance between the null and the positive findings in a purely quantitative way, the explanation in terms of a false positive is the most logical interpretation. However, we argue that other explanations might also be worth considering. The following section addresses these explanations and attempts to draw a general account of our results.

6.2 Comprehensive account of our findings

When reflecting upon a general account that fits our mixed pattern of results, we are challenged by either explaining the scarcity of the positive results or the more frequent null results. If we endorse the original TBRS model's account of switching (as the majority of evidence suggests to do), we need to explain the observed variation of memory performance as a function of our manipulation of the switches (Exp. 6b). In other words, we have to answer the following question: if switch costs between storage and processing cannot explain these results (because there is an equal amount of switches in our experimental manipulations according to the views of the original TBRS model), what does? Alternatively, if we endorse the TBRS2 views of switching (despite obvious support), why did we fail to find such results all along? In the following, we attempt to tackle both challenges.

In a first section, we will begin by assessing an alternative interpretation for the detrimental effect of switching on memory performance, which we observed in Exp. 6b. This explanation based on the role of elaboration is compatible with the conceptualization of the switches of the original TBRS model. In a second section, we will depart from the conceptualization of the switches of the original TBRS model and discuss the possibility that our manipulation of switching did capture, as we hypothesized, switch costs between storage and processing (and the other way around). Here, we will attempt to explain why such findings were not observed in all experiments. In this section, we will consider that the root cause of these mixed findings might be the two maintenance mechanisms posited by the TBRS model, refreshing and rehearsal.

6.2.1 Elaboration

Throughout this dissertation, we have assumed that the subtle memory performance variations that we have observed in one experiment (6b) are caused by our manipulation of the switches. However, as we discussed in the previous section, the body of evidence in favor of our conceptualization of the switches is scarce in the remaining experiments. Therefore, it could be argued that our presumed switch cost result reflects in fact other processes, such as elaboration.

Elaboration refers to deep-processing encoding that is known to enhance long-term memory performance (Craik & Tulving, 1975; Gallo, Meadow, Johnson, & Foster, 2008; Nieznański, 2020). This deep processing is achieved by connecting the newly formed representations with previous knowledge and semantic associations (Craik & Lockhart, 1972; Craik & Tulving, 1975). As we manipulated the structure of the processing task by varying the order of processing steps and free delays, we also varied consequently the response-stimulus interval (RSI) between two distractors. For instance, the processing patterns used in Exp.6 that have 3 and 7 switches (according to our TBRS2 conceptualization) are associated to free delay opportunities of 2400 ms and 1200 ms respectively (in the slow CL conditions) between two processing steps. During these free delays, participants could try for example to visualize or create meaningful sentences from the memoranda they have stored in working memory. Therefore, it could be argued that better memory performance in low switch trials (again according to our TBRS2 conceptualization) found in Exp. 6b arises in fact from the longer RSIs that allow more opportunities to enrich the memory traces.

In summary, the elaboration perspective views the subtle performance variations as the result of a gain from enriched memory traces and not from a switching cost. An interpretation of our positive finding in terms of the role of elaboration is compatible with the conceptualization of the switches of the original TBRS model. The number of switches between storage and processing (and the other way around) is indeed irrelevant according to this account. However, it is worth noting that a specific boosting mechanism could (depending on the magnitude of its effect) potentially lead to unbalance the ratio between the time during which memory traces suffer from decay and the time during which attention is requested for the maintenance activities (i.e. cognitive load).

Although an explanation of our findings based on elaboration seems appealing at first glance, our data do not support this account beyond the results of Experiment 6b. Despite using the same processing patterns, the results of Exp. 3 and 4 did not reveal enhanced memory performance associated to longer RSIs. One could, however, argue that this comparison across experiments is hazardous because we used shorter delays of free time between two distractors in both experiments that revealed null results. For example, in Experiment 6b the free delay opportunities between two distractors were twice as long as those used in Experiment 4 for the same processing patterns (i.e. 2400 ms and 1200 ms for the slow CL condition in Exp. 6b *vs.* 1200 ms and 600 ms in Exp. 4). If elaboration requires long uninterrupted delays of free time to enrich the memory traces, doubling the amount of free time between two processing steps could easily explain the discrepancy in our results. However, there are several reasons based on previous empirical findings that make us doubt this interpretation. Firstly, we know from previous studies that only a modest proportion of participants report actually using spontaneously elaboration strategies (Bailey, Dunlosky, & Kane, 2011; Bartsch & Oberauer, 2021). None of our experiments encouraged to use specific strategies, so there is no reason to believe that elaboration was effectively used by the majority of participants. Secondly and most importantly, even if participants did engage in elaboration, recently conducted experimental work showed no evidence supporting a beneficial effect of elaboration on working memory performance (Bartsch, Loaiza, Jäncke, Oberauer, & Lewis-Peacock, 2019; Bartsch & Oberauer, 2021; Bartsch, Singmann, & Oberauer, 2018).

To conclude, the elaboration account does not provide a satisfactory explanation of our data. It is, however, possible that other mechanisms than elaboration or switching were prompted by the structure of the task. However, for now we fail to find any plausible alternative account. For this reason, we will shift perspective in the next section and assume that our manipulation of the switches did allow to detect a detrimental effect of switching on memory performance. Consequently, this requires to explain why we did not find similar switch costs in all of our experiments.

6.2.2 Refreshing & rehearsal

In this section, we assume (despite obvious support) that the TBRS2 account of the switches reflected accurately, at least in some cases, the executive processes linked to the alternations between storage and processing activities. The following discussion will provide a possible account of our null results based on refreshing and rehearsal which is line with previous work.

A closer examination of Experiment 6a and 6b is particularly relevant in order to understand the determinant factors to observe switch costs between the two main functions of working memory. These two experiments used the same design with the only difference that the task required either keyboard responses (6a) or oral responses (6b). Our results showed that memory performance suffered from a detrimental effect of switching only when oral responses were required. According to the TBRS model, this finding could be explained by the difference in the attentional demands of refreshing and articulatory rehearsal. Refreshing is conceived as an attention demanding process, while articulatory rehearsal is thought to consume only very little attention, if any at all (Camos et al., 2009; Vergauwe et al., 2014). In other words, refreshing competes for the same attentional resource than the execution of the processing task. Articulatory rehearsal on the other hand does not compete with the processing task because it requires no attention after an initial setup period (Naveh-Benjamin & Jonides, 1984). This essentially means that articulatory rehearsal can be used as a cost-free strategy with no trade-off with the processing activity. With this in mind, it is no coincidence that we were not able to detect switch costs in the experiments that did not impede articulatory rehearsal (Exp. 2, 3, 6a). The lack of positive results can indeed be explained by assuming that participants could rehearse the memoranda without any attentional cost, thus potentially neutralizing switch costs associated to memory performance. This account also fits our findings regarding the concurrent task of Exp. 6a. The fact that we did not find switch costs associated to the memory task while such costs were observed on the RTs of the concurrent task could suggest that there was no trade-off between storage and processing activities when the task did not require verbalization.

The difference in the attentional demands of refreshing and rehearsal processes offers an interesting explanation, however, Experiment 4 did not detect statistically significant switch costs on memory performance despite the requirement of oral responses in the concurrent task. Here, the TBRS model can again provide an explanation for this apparent discrepancy in our results. According to the model, rehearsal and refreshing are two independent mechanisms that can be used jointly to maintain verbal information (Camos, 2015; Camos & Barrouillet, 2014; Camos et al., 2009). Experiment 4 used a reading digit processing task and Experiment 6b required from the participants to read the digit *and* make a parity judgment aloud. Therefore, the two concurrent tasks differed not only in the duration of the articulatory suppression but also in their attentional demands. A parity judgment task is likely to involve more central attention than reading digits, which could in turn impede more memory performance. At the same time, the rehearsal activity was also blocked for a longer duration in Exp.6b, which could also lead to poorer recall. Since both mechanisms enable maintenance of verbal information, it is not possible to pinpoint the exact contribution of either articulatory suppression or attentional demand in our results. Nevertheless, it appears that our experimental manipulation was the most effective when both mechanisms were required for the maintenance activities.

In summary, we believe that an explanation of our results in terms of the role of refreshing and rehearsal allows to bring to light a coherent account of our findings. However, it is still too early to draw any definitive conclusions. Despite the fact that the mechanisms responsible for our findings are not entirely elucidated, we believe that our results highlight the importance of implementing variations in the structure of the traditional paradigms used in the literature. Such variations can unveil the fact that there might be other mechanisms at play in working memory that have been neglected so far. We also believe that our findings are a valuable contribution to the field especially because subtle task variations have been overlooked in the past.

6.3 Limits & perspectives

To conclude the present dissertation, we will now address the limitations and the perspectives of our work. Throughout this thesis, we have voiced our concerns about several methodological limits of our experiments. In the following, we will focus on the two most important limitations that we have identified, namely our processing patterns and our conceptualization of the switches. Lastly, we will review future perspectives before turning to our final words regarding this work.

One of the main concerns in our experimental method is the processing patterns that might have tapped into other aspects than just task-switching processes. This was particularly clear in our first experiments (Exp. 2-5) in which several confounds were introduced, making thus interpretations of the results uneasy. In detail, Exp. 2 revealed that the number of switches was confounded with the cognitive load. In Experiments 3-5 we remedied this confound by using six processing schedules that controlled for the duration of the retention intervals, memory load, and the amount and nature of distractors in a trial. However, the numerous processing schedules that we used tapped perhaps into other mechanisms such as task predictability. The inherent limitations of our experimental material were also clear when we attempted to investigate other processes than switching. In this regard, four experiments examined the hypothesis that an immediate free delay after the presentation of a storage item is beneficial to working memory performance. Our results confirmed this observation only in Experiment 4. In the working memory literature this finding is typically associated to a consolidation account which assumes that the immediate free delay in question will strengthen the memory traces, which leads in turn to enhanced memory performance. As discussed in Chapter 2 and 5, we suggested that the beneficial effect of the free delay could also be explained by alternative accounts, such as task preparation or discrepancies in the cumulative refreshment schedule (following the TBRS2 predictions). Unfortunately, we were unable to disentangle the different accounts from each other. Interpretation would have been facilitated given a more successful outcome because the confounds would have been less plausible.

In sum, it clearly appears that the used processing patterns (especially when numerous patterns were used in one experiment) have made interpretation of our results uneasy. On the other hand, we believe that these patterns were useful because they allowed to bring insight to macro and micro task-structure variations in the complex span task paradigm. The method let us discover that these aspects should not be neglected, as they can influence working memory performance. In a more general way, the processing patterns allowed also to test a benchmark (Benchmark 5.2.4.), namely the cognitive load effect as conceptualized by the TBRS model. An effect considered as a benchmark means that the finding is robust and should generalize across materials and methodological variations (Oberauer et al., 2018b). The study of structural variations of complex span task allowed thus to examine how subtle changes in material and methodology might affect cognitive load which is in itself valuable to the community.

Another important limitation of our work is related to our conceptualization of the switches. Here, we suggested two different views how switches could be conceived (TBRS *vs.* TBRS2). However, the experimental method did not allow to address whether all switches were equal. In other words, we are unable to determine whether the switches occurring from storage or from processing contribute differently to our findings. On the other hand, we believe that our approach was a good starting point to explore specifically switches occurring between the two main functions of working memory. More importantly, our approach allowed to examine switching in a complex span task without completely altering the paradigm so that generalization of our findings to the most commonly used working memory task could be easily made. Unfortunately, at this stage we are unable to draw definitive conclusions regarding the exact mechanisms that might have contributed to our findings. Given the surge of interest in the consolidation effect and in switching between different sources of information (i.e. external *vs.* internal), we remain hopeful that in due course these phenomenons will be better understood. With this in mind, the following discussion offers few perspectives related to our work that could perhaps help further the field.

We believe that an interesting avenue of research relates to clarifying the role of articulatory rehearsal in our findings. As discussed elsewhere, in one experiment when our task required verbal responses, we were able to observe switching costs associated to working memory performance. Regarding this finding, the idea that verbalization plays a role in task switching is not a new one (e.g. Bryck & Mayr, 2005; Liefooghe et al., 2005; Miyake et al., 2004; Saeki & Saito, 2004, 2009), but because our task required oral responses and not articulatory suppression per se, it would be interesting to see if the observed cost could be increased if a similar experimental design was carried out using articulatory suppression. The use of visual material that varies on their degree of verbalizability could be another option to study this question. For instance, the study of Arslan, Broc, and Mathy (2020) used in span tasks four different classes of visual stimuli: (1) easily verbalizable images that bear similitude with meaningful symbols; (2) non-verbalizable images with no resemblance to meaningful symbols that are thus unlikely to be recalled with verbal cues; (3) images of astronomical objects (e.g. galaxies) that look like everyday objects (e.g. chair) that could perhaps be recalled by semantic association; (4) images of astronomical objects that are non-verbalizable, and that should be recalled in all likelihood with visual cues only. With the use of such a material, we could imagine a modulation of the switch costs as a function of the number of switches required by the task and the degree of verbalizability of the material. The advantage of such a material is that it would require from the participant essentially the use of the refreshing maintenance mechanism according to the TBRS model, which should render the detection of switch costs more likely.

Our results also showed that cognitive load of the task is an important aspect to take into consideration when exploring switch costs in complex span tasks. We discussed in Chapter 4 that it seems that an optimal cognitive load condition is needed to observe switching costs. An intuitive interpretation is that when the cognitive load of the task is too low, a switch might not tax performance enough to be noticeable. Despite of the putative switch cost, it is possible that participants could manage to compensate for the cost and reach their span level because they had sufficient time remaining to do so. In other situations involving a task that is highly demanding, participants might choose to prioritize either the memory or the concurrent task. Such a strategic choice would essentially mean that participants do not switch regularly between storage and processing activities. Thus, the detection of switch costs could be compromised. Closely related to this point, participants could also make speed-accuracy adjustments within the concurrent task, which impacts our ability to detect switch costs. Participants could slow down the execution of the concurrent task involving numerous switches in order to achieve more accurate performance. Such a pattern of result was indeed observed in Experiment 6a. Individual differences in working memory are likely to aggravate our difficulty to find the “optimal cognitive load” to capture potential switch costs. Given that working memory capacity differs from one participant to another, the degree of performance impairment resulting from a cognitive load manipulation is bound to be conditioned by the capacity of each individual. For instance, what constitutes a low cognitive load for one individual might correspond for another participant to a high load (Doherty et al., 2019). Therefore, if we want to capture switch costs between storage and processing activities, it is primordial to ensure that participants are actually able to switch between the two activities. Several steps could be considered to remedy these issues. First, the use of tasks calibrated to the participants’ individual capacity might be an interesting solution to consider. We adopted this strategy in our last experiments examining the consolidation effect (Exp. 7-8), and we believe that such a method could be particularly interesting for the study of switching effects in working memory performance as well. This could perhaps allow to find the optimal cognitive load for each individual more easily. Second, to avoid the difficulty related to the interpretation of speed-accuracy biases, we could use a component measure of both accuracy and latency scores of the concurrent task as we have discussed in section 6.1.1.

A complementary approach to tackle switching in complex span tasks would be to take advantage of the rich task-switching literature. One option could be to merge complex span tasks with existing switching paradigms to follow the rationale of the task-span procedure Logan (2004). In our case, we could envision a complex span task using an alternative-runs procedure with and without a memory load (i.e. AABBAABB). We think this is the best next option that is perhaps worth developing with larger populations (e.g. children). In a similar vein, a promising avenue to disentangle consolidation from other accounts could be to merge the complex span task with existing paradigms that are known to efficiently examine the effect of task preparation on performance. For instance, this could involve examining the effect of extensive practice on the consolidation effect. The rationale here is that if the consolidation effect dissipates with practice, then we could argue that it is unlikely that specific encoding mechanism are the cause of the effect in question. Computational modelling could also be an invaluable approach to examine whether discrepancies in the cumulative refreshment schedule can possibly account, at least in part, for the observed effect. For this purpose the TBRS2 is a helpful tool, as it allows to easily assess with precision the subtle discrepancies in performance which could be attributed to the refreshing schedule.

Finally, it might be worth considering that traditional behavioral measures are perhaps not sensitive enough to detect such transient switch costs. Within the task-switching literature several other markers of switching behavior have been identified. For instance, task-set preparation has been linked to an anticipatory orientation of gaze to the location of the next task and a positive polarity brain potential over the posterior cortex (Longman, Elchlepp, Monsell, & Lavric, 2021). The use of (concurrent) EEG and eye-tracking might perhaps shed light on whether we can find similar indicators of task switching in a complex span task and thus allow more sensitive measures of this seemingly fleeing phenomenon.

6.4 Conclusion

For the purpose of the present dissertation, we created new complex span tasks that manipulated either the global structure of a task or more subtle variations of the processing episodes. Our findings revealed that changes in the task structure should not be overlooked, as they can easily influence working memory performance and introduce confounds. The question of what factors specifically drives these variations of performance is, however, not completely elucidated, although some of our results suggest that switch costs could be in part responsible.

Although our findings suggest that such costs are likely to tax working memory performance, it is too early to draw any definitive conclusions in light of the little evidence in their favor. Beyond the question of a putative cost of the switch in working memory complex span tasks, several questions also remain open-ended such as when exactly do we switch, what is their root cause, and are all types of switch equal?

The existence of switch costs taxing working memory performance would be compatible with the time-based resource-sharing model which considers switching as a key ability of working memory functioning. In so far, the model had not addressed explicitly whether switching between storage and processing activities (and the other way around) in working memory is associated to a specific cost or not. Our findings point to the conclusion, however, that the TBRS model in its current implementation remains the most parsimonious way to account for performance in complex span tasks, regardless of subtle variations of performance that can be introduced when manipulating the structure of the task.

References

Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, *3*(301).

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological bulletin*, *131*(1), 30.

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, Speech, and Hearing Services in Schools*, *49*(3), 340–355.

Akaike, H., Petrov, B. N., & Csaki, F. (1973). *Second international symposium on information theory.* Akadémiai Kiadó, Budapest.

Allen, R., Baddeley, A., & Hitch, G. (2014). Evidence for two attentional components in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1499.

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*(1), 64–88.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks.

Altmann, E. M. (2007). Comparing switch costs: Alternating runs and explicit cuing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 475.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological science*, *13*(1), 27–33.

Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive psychology*, *30*(3), 221–256.

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychological science*, *15*(9), 610–615.

Arslan, S., Broc, L., & Mathy, F. (2020). Lower verbalizability of visual stimuli modulates differences in estimates of working memory capacity between children with and without developmental language disorders. *Autism & Developmental Language Impairments*, *5*, 2396941520945519.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, *225*(2), 82–91.

Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 5–28.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, *4*(11), 417–423.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of experimental psychology: General*, *130*(4), 641.

Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology*, *11*(2), 187.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, *105*(1), 158.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Recent advances in learning and motivation* (p. 647-667). New York, NY: Academic Press.

Baddeley, A., & Hitch, G. (2007). Working memory: Past, present... and future. *The cognitive neuroscience of working memory*, 1–20.

Baddeley, A., Hitch, G., & Allen, R. (n.d.). (R. Logie, V. Camos, & N. Cowan, Eds.). Oxford University Press.

Baddeley, A., Hitch, G., & Allen, R. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, *36*(2), 233–252.

Baddeley, A., & Logie, R. (1999). Working memory: The multiple-component model.

Baddeley, A., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *The Quarterly Journal of Experimental Psychology Section A*, *38*(4), 603–618.

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *14*, 575-589.

Bailey, H. (2012). Computer-paced versus experimenter-paced working memory span tasks: Are they equally reliable and valid? *Learning and Individual Differences*, *22*(6), 875–881.

Bailey, H., Dunlosky, J., & Kane, M. J. (2011). Contribution of strategy use to performance on complex and simple span tasks. *Memory & cognition*, *39*(3), 447–461.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*, 83-100.

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*, 570-585.

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

Barrouillet, P., & Camos, V. (2007). The time-based ressource-sharing model of working memory. In *The cognitive neuroscience of working memory* (pp. 455, 59–80.).

Barrouillet, P., & Camos, V. (2009). Interference: unique source of forgetting in working memory? *Trends in Cognitive Sciences*, *13*(4), 145–146.

Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*, *21*(6), 413–419.

Barrouillet, P., & Camos, V. (2014a). On the proper reading of the TBRS model: reply to Oberauer and Lewandowsky (2014). *Front. Psychol.*, *5*.

Barrouillet, P., & Camos, V. (2014b). *Working memory: Loss and reconstruction*. Psychology Press.

Barrouillet, P., & Camos, V. (2020). *The time-based ressource-sharing model of working memory.* (R. Logie, V. Camos, & N. Cowan, Eds.). Oxford University Press.

Barrouillet, P., De Paepe, A., & Langerock, N. (2012). Time causes forgetting from working memory. *Psychon Bull Rev*, *19*, 87–92.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: a time-based resource-sharing model account. *Developmental Psychology*, *45*, 477-490.

Barrouillet, P., Lepine, R., & Camos, V. (2008). Is the influence of working memory capacity on high-level cognition mediated by complexity or resource-dependent elementary processes? *Psychonomic Bulletin & Review*, *15*, 528–534.

Barrouillet, P., Portrat, S., & Camos, V. (2011). On the law relating processing to storage in working memory. *Psychological Review*, 175-192.

Bartsch, L. M., Loaiza, V. M., Jäncke, L., Oberauer, K., & Lewis-Peacock, J. A. (2019). Dissociating refreshing and elaboration and their impacts on memory. *NeuroImage*, *199*, 585–597.

Bartsch, L. M., & Oberauer, K. (2021). The effects of elaboration on working memory and long-term memory across age. *Journal of Memory and Language*, *118*, 104215.

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*(5), 796–808.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

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.

Belletier, C., Camos, V., & Barrouillet, P. (2020). Is the cognitive system much more robust than anticipated? dual-task costs and residuals in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Bernardin, S., Portrat, S., & Barrouillet, P. (2006). Impact de la nature et du rythme de réalisation de la tâche secondaire dans les tâches d'empan de mémoire de travail. *Année psy*, *106*, 23.

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009, July). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition*, *37*(5), 689–713.

Blalock, L. D. (2015). Stimulus familiarity improves consolidation of visual working memory representations. *Atten Percept Psychophys, 77*, 1143–1158.

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Broadbent, D. E. (1975). The magic number seven after fifteen years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long-term memory.* New York: Wiley.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539-576.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*, 127-181.

Bryck, R. L., & Mayr, U. (2005). On the role of verbalization during task set selection: Switching or serial order control? *Memory & Cognition*, *33*(4), 611–623.

Bryck, R. L., & Mayr, U. (2008). Task selection cost asymmetry without task switching. *Psychonomic Bulletin & Review*, *15*(1), 128–134.

Burgess, N., & Hitch, G. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551-581.

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of memory and language*, *31*(4), 429–460.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*(4), 627–652.

Burgess, P. W., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in cognitive sciences*, *11*(7), 290–298.

Burgess, P. W., Gilbert, S. J., & Dumontheil, I. (2007). Function and localization within rostral prefrontal cortex (area 10). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 887–899.

Burnham, K. P., & Anderson, D. R. (2002). A practical information-theoretic approach. *Model selection and multimodel inference, 2nd ed. Springer, New York*, *2*.

Byrnes, D. L., & Wingfield, A. (1979). Retrieval time and decay of information in dichotic memory. *Perceptual and Motor Skills*, *48*(3), 831–839.

Camos, V. (2015). Storing verbal information in working memory. *Current Directions in Psychological Science*, *24*(6), 440–445.

Camos, V. (2017). Domain-specific versus domain-general maintenance in working memory: Reconciliation within the time-based resource sharing model. *Psychology of learning and motivation*, *67*, 135–171.

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. *Frontiers in human neuroscience*, *8*, 900.

Camos, V., Johnson, M. R., Loaiza, V. M., Portrat, S., Souza, A. S., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 19–32.

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*, 457–469.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of experimental child psychology*, *33*(3), 386–404.

Cauchi, C. (2016). *Proposition de tâches concurrentes variables sur le plan rythmique dans le modèle du partage des ressources basé sur le temps.* (Université Nice Sophia-Antipolis, France ed.). Mémoire de Master 1 non publié.

Cohen, J. (1994). The earth is round (p<. 05). *American psychologist*, *49*(12), 997.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of verbal learning and verbal behavior*, *15*(1), 17–31.

Colom, R., Flores-Mendoza, C., & Rebollo, I. (2003). Working memory and intelligence. *Personality and Individual Differences*, *34*(1), 33–39.

Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: a reanalysis of key studies. *Memory & Cognition*, *34*, 158-171.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, *55*, 75-84.

Conway, A. R., 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*, 769-786.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*, 547-552.

Corbin, L., Moissenet, A., & Camos, V. (2012). Fonctionnement de la mémoire de travail chez des enfants présentant des difficultés scolaires. *Développements*, *11*, 5.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and l2 learners. *Journal of memory and Language*, *50*(4), 491–511.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control*, *20*, 506.

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. (2008). What are the differences between long-term, short-term, and working memory? *Progress in brain research*, *169*, 323–338.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educ Psychol Rev*, *26*(2), 197–223.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychon Bull Rev*, *24*(4), 1158–1170.

Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language*, *31*, 1-17.

Cowan, N., Keller, T. A., Hulme, C., Roodenrys, S., McDougall, S., & Rack, J. (1994). Verbal memory span in children: Speech timing clues to the mechanisms underlying age and word length effects. *Journal of Memory and Language*, *33*(2), 234–250.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671–684.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general*, *104*(3), 268.

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

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic bulletin & review*, *3*(4), 422–433.

Darling, S., Parker, M.-J., Goodall, K. E., Havelka, J., & Allen, R. J. (2014). Visuospatial bootstrapping: Implicit binding of verbal working memory to visuospatial representations in children and adults. *Journal of Experimental Child Psychology*, *119*, 112–119.

Derks, P. L. (1974). The length-difficulty relation in immediate serial recall. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 335–354.

De Schrijver, S., & Barrouillet, P. (2017). Consolidation and restoration of memory traces in working memory. *Psychon Bull Rev*, *24*, 1651–1657.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, *6*(3), 274–290.

Doherty, J. M., Belletier, C., Rhodes, S., Jaroslawska, A., Barrouillet, P., Camos, V., . . . Logie, R. H. (2019). Dual-task costs in working memory: An adversarial collaboration. *Journal of experimental psychology: learning, memory, and cognition*, *45*(9), 1529.

Draheim, C., Hicks, K. L., & Engle, R. W. (2016). Combining reaction time and accuracy: The relationship between working memory capacity and task switching as a case example. *Perspectives on Psychological Science*, *11*(1), 133–155.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245.

Ebbinghaus, H. (1885/1913). Über das gedächtnis: Untersuchen zur experimentellen psychologie [memory: A contribution to experimental psychology](ha ruger & h. bussenius, trans.). *Leipzig, Germany: Duncker & Humblot.(Original work published 1885)*.

Ellefson, M. R., Shapiro, L. R., & Chater, N. (2006). Asymmetrical switch costs in children. *Cognitive Development*, *21*(2), 108–130.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24,* 143-188.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48,* 148 - 168.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19-23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190–193.

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.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory for letter strings. *Journal of Verbal Learning & Verbal Behavior*, *12*(4), 360-372.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*, 59-79.

Farrell, S., & Lewandowsky, S. (2010). Computational models as aids to better reasoning in psychology. *Current Directions in Psychological Science*, *19*(5), 329–335.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using g* power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149–1160.

Filippi, R., Karaminis, T., & Thomas, M. S. (2014). Language switching in bilingual production: Empirical data and computational modelling. *Bilingualism: Language and Cognition*, *17*(2), 294–315.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition*, *43*(2), 226–236.

Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, *37*, 581-590.

Gallistel, C. R. (2009). The importance of proving the null. *Psychological review*, *116*(2), 439.

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language*, *58*(4), 1095–1111.

Garavan, H. (1998). Serial attention within working memory. *Memory & cognition*, *26*(2), 263–276.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*(11), 410-419.

Gathercole, S. E., & Adams, A.-M. (1994). Children s phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of memory and language*, *33*(5), 672–688.

Gauvrit, N., & Mathy, F. (2018). Mathematical transcription of the 'time-based resource sharing'theory of working memory. *British Journal of Mathematical and Statistical Psychology*, *71*, 146–166.

Gignac, G. E. (2015). The magical numbers 7 and 4 are resistant to the Flynn effect: No evidence for increases in forward or backward recall across 85 years of data. *Intelligence*, *48*, 85-95.

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General*, *129*(3), 308.

Goschke, T. (2000). Intentional reconfiguration and j-ti involuntary persistence in task set switching. *Control of cognitive processes: Attention and performance XVIII*, *18*, 331.

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science*, *16*, 70-76.

Halford, G. S., Mayberry, M., & Bain, J. (1988). Set-size effects in primary memory: An aged-related capacity limitation. *Memory & Cognition*, *16*(480-487).

Halford, G. S., Wilson, W. H., & Phillips, W. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral and Brain Sciences*, *21*, 803-831.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. *Brain Mechanisms and Learning*, 37-46.

Hedge, C., & Leonards, U. (2013). Using eye movements to explore switch costs in working memory. *Journal of Vision*, *13*(4), 18–18.

Henseler, I., Krüger, S., Dechent, P., & Gruber, O. (2011). A gateway system in rostral pfc? evidence from biasing attention to perceptual information and internal representations. *Neuroimage*, *56*(3), 1666–1676.

Henson, R. (1998). Short-term memory for serial order: The start-end model. *Cognitive psychology*, *36*(2), 73–137.

Henson, R. (1999). Positional information in short-term memory: Relative or absolute?. *Memory & Cognition*, *27*, 915-927.

Hübner, M., Kluwe, R. H., Luna-Rodriguez, A., & Peters, A. (2004). Response selection difficulty and asymmetrical costs of switching between tasks and stimuli: no evidence for an exogenous component of task-set reconfiguration. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(6), 1043.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior research methods*, *46*(3), 702–721.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory. *Journal of Experimental Child Psychology*, *38*, 241-253.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*(5), 497–514.

Jacobs, J. (1887). Experiments on Prehension. *Mind*, *12*(45), 75-79.

James, W. (1890). The principles of psychology (vol. 1) new york: Henry holt & co.

Janczyk, M., Wienrich, C., & Kunde, W. (2008). On the costs of refocusing items in working memory: A matter of inhibition or decay? *Memory*, *16*(4), 374–385.

Jarrold, C., Tam, H., Baddeley, A., & Harvey, C. (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.

Jersild, A. T. (1927). Mental set and shift. *Archives of psychology*.

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive psychology*, *36*(2), 138–202.

Jones, D. (1993). Objects, streams, and threads of auditory attention.

Jones, D., Hughes, R., & Macken, W. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology*, *60*(4), 505–511.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193-224.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on ackerman, beier, and boyle (2005). *Psychological Bulletin*, *131*, 66–71.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189.

Kanerva, K., & Kalakoski, V. (2016). The predictive utility of a working memory span task depends on processing demand and the cognitive task: The predictive utility of WM span task. *Applied Cognitive Psychology*, *30*, 681–690.

Kelley, M. R., Neath, I., & Surprenant, A. M. (2013). Three more semantic serial position functions and a simple explanation. *Memory & Cognition*, *41*(4), 600–610.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of verbal learning and verbal behavior*, *1*(3), 153–161.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1998). *Insights into working memory from the perspective of the epic architecture for modeling skilled perceptual-motor and cognitive human performance.* (Tech. Rep.).

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—a review. *Psychological bulletin*, *136*(5), 849.

Kiesel, A., Wendt, M., & Peters, A. (2007). Task switching: On the origin of response congruency effects. *Psychological Research*, *71*(2), 117–125.

Koch, I. (2003). The role of external cues for endogenous advance reconfiguration in task switching. *Psychonomic bulletin & review*, *10*(2), 488–492.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433.

Langerock, N., Vergauwe, E., & Barrouillet, P. (2014). The maintenance of cross-domain associations in the episodic buffer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 1096.

Laurent, L., Millot, J.-L., Andrieu, P., Camos, V., Floccia, C., & Mathy, F. (2016). Inner speech sustains predictable task switching: direct evidence in adults. *Journal of Cognitive Psychology*, *28*(5), 585–592.

Lechner, H. A., Squire, L. R., & Byrne, J. H. (1999). 100 years of consolidation—remembering müller and pilzecker. *Learning & Memory*, *6*(2), 77–87.

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.

Lemaire, B., & Portrat, S. (2018). A computational model of working memory integrating time-based decay and interference. *Frontiers in psychology*, *9*, 416.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, *11*, 771-790.

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. SAGE publications.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25-57.

Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. *Psychological Review*, *122*(4), 674.

Lewandowsky, S., Oberauer, K., Yang, L. X., & Ecker, U. K. (2010). A working memory test battery for matlab. *Behavior Research Methods*, *42*, 571-585.

Liefooghe, B., Barrouillet, P., Vandierendonck, A., & Camos, V. (2008). Working memory costs of task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 478–494.

Liefooghe, B., Vandierendonck, A., Muyllaert, I., Verbruggen, F., & Vanneste, W. (2005). The phonological loop in task alternation and task repetition. *Memory*, *13*(5), 550–560.

Loaiza, V. M., & Camos, V. (2016). Does controlling for temporal parameters change the levels-of-processing effect in working memory? *Advances in cognitive psychology*, *12*(1), 2.

Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science*, *12*(2), 45–48.

Logan, G. D. (2004). Working memory, task switching, and executive control in the task span procedure. *Journal of Experimental Psychology: General*, *133*, 218-236.

Logan, G. D. (2006). Out with the old, in with the new: More valid measures of switch cost and retrieval time in the task span procedure. *Psychonomic Bulletin & Review*, *13*(1), 139–144.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological review*, *108*(2), 393.

Logie, R. H. (1986). Visuo-spatial processing in working memory. *The Quarterly Journal of Experimental Psychology Section A*, *38*(2), 229–247.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In *Advances in psychology* (Vol. 80, pp. 105–115). Elsevier.

Longman, C. S., Elchlepp, H., Monsell, S., & Lavric, A. (2021). Serial or parallel proactive control of components of task-set? a task-switching investigation with concurrent eeg and eye-tracking. *Neuropsychologia*, 107984.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after spearman's (1904)"'general intelligence,'objectively determined and measured". *Journal of personality and social psychology*, *86*(1), 96.

Lucidi, A., Langerock, N., Hoareau, V., Lemaire, B., Camos, V., & Barrouillet, P. (2016, February). Working memory still needs verbal rehearsal. *Mem Cogn*, *44*(2), 197–206.

Lucidi, A., Loaiza, V., Camos, V., & Barrouillet, P. (2014). Assessing working memory capacity through time constrained elementary activities. *The Journal of General Psychology*, *141*, 98–112.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Lépine, R. l., Barrouillet, P., & Camos, V. (2005, February). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review*, *12*(1), 165–170.

Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and complex working memory tasks allow similar benefits of information compression. *Journal of Cognition*, *1*(1).

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*, 346-362.

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 362.

McCabe, D. P. (2010). The influence of complex working memory span task administration methods on prediction of higher level cognition and metacognitive control of response times. *Memory & Cognition, 38*(7), 868–882.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological review*, *39*(4), 352.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, memory, and cognition*, *22*(6), 1423.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning & Verbal Behavior*, *2*, 1-21.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*, *40*(1), 25–40.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81-97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J.-c. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta psychologica*, *115*(2-3), 123–142.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49-100.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*, 134 - 140.

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 493.

Monsell, S., et al. (1996). Control of mental processes. *Unsolved mysteries of the mind: Tutorial essays in cognition*, 93–148.

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological research*, *63*(3), 250–264.

Morey, C. C., & Cowan, N. (2018). Can we distinguish three maintenance processes in working memory?: Maintenance processes in working memory. *Ann. N.Y. Acad. Sci.*, *1424*(1), 45–51.

Morey, R. D., Romeijn, J.-W., & Rouder, J. N. (2016). The philosophy of bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, *72*, 6–18.

Morris, N., & Jones, D. M. (1990). Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, *81*, 111-121.

Mueller, S. T., & Krawitz, A. (2009). Reconsidering the two-second decay hypothesis in verbal working memory. *Journal of Mathematical Psychology*, *53*(1), 14–25.

Murdock, B. B. (1960). The distinctiveness of stimuli. *Psychological Review*, *67*, 16-31.

Murray, D. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*(4p1), 679.

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(3), 369.

Nicholson, R., Karayanidis, F., Davies, A., & Michie, P. T. (2006). Components of task-set reconfiguration: Differential effects of 'switch-to'and 'switch-away'cues. *Brain research*, *1121*(1), 160–176.

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.

Nieznański, M. (2020). Levels-of-processing effects on context and target recollection for words and pictures. *Acta Psychologica*, *209*, 103127.

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1–18). Springer.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 411.

Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition*, *37*, 346-357.

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the time-based resource-sharing theory. *Psychonomic Bulletin & Review*, *18*, 10-45.

Oberauer, K., & Lewandowsky, S. (2014). Further evidence against decay in working memory. *Journal of Memory and Language*, *73*, 15–30.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., ... others (2018a). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., ... Ward, G. (2018b). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: an interference model of complex span. *Psychonomic Bulletin & Review*, *19*, 779-819.

Page, M. P., & Norris, D. (1998). The primacy model: a new model of immediate serial recall. *Psychological Review*, *105*, 761-781.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. *Psychological bulletin*, *116*(2), 220.

Pashler, H. (1999). *The psychology of attention*. MIT press.

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*, 421-425.

Plancher, G., & Barrouillet, P. (2013). Forgetting from working memory: Does novelty encoding matter? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 110–125.

Portrat, S., Camos, V., & Barrouillet, P. (2009). Working memory in children: A time-constrained functioning similar to adults. *Journal of Experimental Child Psychology*, *102*(3), 368–374.

Portrat, S., & Lemaire, B. (2015). Is attentional refreshing in working memory sequential? a computational modeling approach. *Cogn Comput*, *7*(3), 333–345.

Raven, J. C. (1962). *Advanced progressive matrices: Sets i and II*. HK Lewis.

Ravizza, S. M., & Carter, C. S. (2008). Shifting set about task switching: Behavioral and neural evidence for distinct forms of cognitive flexibility. *Neuropsychologia*, *46*(12), 2924–2935.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex*, *43*(1), 135–145.

Rhodes, S., Jaroslawska, A. J., Doherty, J. M., Belletier, C., Naveh-Benjamin, M., Cowan, N., ... 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*(7), 1204.

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.

Ricker, T. J., & Hardman, 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., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory: An overview of working memory consolidation. *Ann. N.Y. Acad. Sci.*, *1424*(1), 8–18.

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).

Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, *69*(10), 1969–1995.

Ricker, T. J., Vergauwe, E., Hinrichs, G. A., Blume, C. L., & Cowan, N. (2015). No recovery of memory when cognitive load is decreased. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 872.

Risse, S., & Oberauer, K. (2010). Selection of objects and tasks in working memory. *Quarterly Journal of Experimental Psychology*, *63*(4), 784–804.

Robbins, T., Anderson, E., Barker, D., Bradley, A., Fearnyhough, C., Henson, R., ... Baddeley, A. (1996). Working memory in chess. *Memory & cognition*, *24*(1), 83–93.

Roebuck, H., Freigang, C., & Barry, J. G. (2016). Continuous performance tasks: Not just about sustaining attention. *Journal of speech, language, and hearing research*, *59*(3), 501–510.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of experimental psychology: General*, *124*(2), 207.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1477.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of experimental psychology: human perception and performance*, *27*(4), 763.

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human perception and performance*, *27*(6), 1404.

Saeki, E., & Saito, S. (2004). Effect of articulatory suppression on task-switching performance: Implications for models of working memory. *Memory*, *12*(3), 257–271.

Saeki, E., & Saito, S. (2009). Verbal representation in task order control: An examination with transition and task cues in random task switching. *Memory & Cognition*, *37*(7), 1040–1050.

Salame, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of verbal learning and verbal behavior*, *21*(2), 150–164.

Salamé, P., & Baddeley, A. (1986). Phonological factors in stm: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society*, *24*(4), 263–265.

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: neuromagnetic evidence. *Journal of cognitive neuroscience*, *5*(3), 363–370.

Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: the gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1089.

Schneider, D. W., & Logan, G. D. (2005). Modeling task switching without switching tasks: a short-term priming account of explicitly cued performance. *Journal of Experimental Psychology: General*, *134*(3), 343.

Schneider, D. W., & Logan, G. D. (2014). Tasks, task sets, and the mapping between them. *Task switching and cognitive control*, 27–44.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of experimental psychology: General*, *125*(1), 4.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly journal of experimental psychology*, *22*(2), 261–273.

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20–31.

Souza, A. S., & Vergauwe, E. (2018). Unravelling the intersections between consolidation, refreshing, and removal. *Annals of the New York Academy of Sciences*, *1424*(1), 5–7.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *The American Journal of Psychology*, 669–679.

Stevanovski, B., & Jolicœur, P. (2007). Visual short-term memory: Central capacity limitations in short-term consolidation. *Visual Cognition*, *15*(5), 532–563.

Stone, J. M., & Towse, J. (2015). A working memory test battery: Java-based collection of seven working memory tasks. *Journal of Open Research Software*, *3*.

Strobach, T., Liepelt, R., Schubert, T., & Kiesel, A. (2012). Task switching: effects of practice on switch and mixing costs. *Psychological research*, *76*(1), 74–83.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Sun, R. (2008). *The cambridge handbook of computational psychology*. Cambridge University Press.

Surprenant, A. M., & Neath, I. (2009). Interactions between short-term and long-term memory in the verbal domain. In A. Thorn & M. P. (Eds.) (Eds.), (p. 16-43). Hove, UK: Psychology Press.

Szmalec, A., Vandierendonck, A., & Kemps, E. (2005). Response selection involves executive control: Evidence from the selective interference paradigm. *Memory & Cognition*, *33*(3), 531–541.

Tekin, E., & Roediger III, H. L. (2020). Reactivity of judgments of learning in a levels-of-processing paradigm. *Zeitschrift für Psychologie*.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). Revisiting the attentional demands of rehearsal in working-memory tasks. *Journal of Memory and Language*, *105*, 1–18.

Towse, J. N., & Hitch, G. (1995). *The Quarterly Journal of Experimental Psychology*, *48*(1), 108–124.

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.

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.

Tulving, E., et al. (1972). Episodic and semantic memory. *Organization of memory*, *1*, 381–403.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological review*, *80*(5), 352.

Turner, M. L., & Engle, R. W. (1989, April). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive psychology*, *5*(3), 275–287.

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68-80.

Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104.

Unsworth, N., & Engle, R. W. (2007b). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038-1066.

Unsworth, N., Heitz, R. P., Schrock, J., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*, 498–505.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: a latent-variable analysis of the relationship between processing and storage. *Memory*, *17*, 635-654.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*, 392–406.

van den Bosch, R. J., Rombouts, R. P., & van Asma, M. J. (1996). What determines continuous performance task performance? *Schizophrenia Bulletin*, *22*(4), 643–651.

Vandierendonck, A., Christiaens, E., & Liefooghe, B. (2008). On the representation of task information in task switching: Evidence from task and dimension switching. *Memory & Cognition*, *36*(7), 1248–1261.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: interplay of reconfiguration and interference control. *Psychological bulletin*, *136*(4), 601.

Vandierendonck, A., Szmalec, A., Deschuyteneer, M., & Depoorter, A. (2007). Towards a multicomponent view of executive control: The case of response selection. *The cognitive neuroscience of working memory*, 247–259.

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.

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological science*, *21*(3), 384–390.

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.

Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval rate may describe many cognitive procedures. *Frontiers in Human Neuroscience*, *8*, 126.

Vergauwe, E., & Cowan, N. (2015). Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychonomic Bulletin & Review*, *22*(4), 1001–1006.

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.

Vergauwe, E., Ricker, T. J., Langerock, N., & Cowan, N. (2019). What do people typically do between list items? The nature of attention-based mnemonic activities depends on task context. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 779–794.

Verschooren, S., Liefooghe, B., Brass, M., & Pourtois, G. (2019). Attentional flexibility is imbalanced: Asymmetric cost for switches between external and internal attention. *Journal of experimental psychology: human perception and performance*, *45*(10), 1399.

Verschooren, S., Pourtois, G., & Egner, T. (2020). More efficient shielding for internal than external attention? evidence from asymmetrical switch costs. *Journal of experimental psychology: human perception and performance*.

Verschooren, S., Schindler, S., De Raedt, R., & Pourtois, G. (2019). Switching attention from internal to external information processing: A review of the literature and empirical support of the resource sharing account. *Psychonomic bulletin & review*, *26*(2), 468–490.

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 and Performance*, *32*(6), 1436.

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., . . . others (2018). Bayesian inference for psychology. part i: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review*, *25*(1), 35–57.

Wang, T., Ren, X., Li, X., & Schweizer, K. (2015). The modeling of temporary storage and its effect on fluid intelligence: Evidence from both brown–peterson and complex span tasks. *Intelligence*, *49*, 84–93.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, *92*(4), 885–896.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive psychology*, *46*(4), 361–413.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological review*, *72*(2), 89.

Wechsler, D. (2011). *Wais-iv nouvelle version de l'échelle d'intelligence de wechsler pour adultes.* Montreuil: ECPA Pearson.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological research*, *63*(3-4), 212–233.

Yeung, N., & Monsell, S. (2003a). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 919.

Yeung, N., & Monsell, S. (2003b). Switching between tasks of unequal familiarity: the role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human perception and performance*, *29*(2), 455.

List of Figures

List of Tables