

WORKING MEMORY CAPACITY: CONCURRENT SUBTASKS NEED NOT INTERFERE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
NEUROSCIENCE

By
Gülsüm Özge Şengil
October 2022

WORKING MEMORY CAPACITY: CONCURRENT SUBTASKS
NEED NOT INTERFERE

By Gülsüm Özge Şengil

October 2022

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Ausaf Ahmed Farooqui(Advisor)

Asst. Prof. Miri Besken

Asst. Prof. Buse Merve Ürgen

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

WORKING MEMORY CAPACITY: CONCURRENT SUBTASKS NEED NOT INTERFERE

Gülsüm Özge Şengil

M.S. in Neuroscience

Advisor: Asst. Prof. Ausaf Ahmed Farooqui

October 2022

Any extended task episode is subsumed by goal-directed programs that hierarchically control its execution. We investigated the relationship between working memory capacity and the control instantiated by such hierarchical task entities across four experiments. In a new extended task consisting of subtask A and subtask B, participants first memorized the orientation of subtask A lines (let's call this event mA), then memorized subtask B lines (mB), then recalled these B lines (rB), and finally recalled A lines (rA). The task structure was: mA-mB-rB-rA. Subtask A lines were thus held in mind during the execution of subtask B. Even though participants had to remember the orientation of lines in both cases, increased WM load of lines A only affected performance on subtask A and did not affect the performance on subtask B. In Experiment 2, four trials of Exp1 were organized into a complex 4-part task with the added condition that A lines of a part be recalled not in that part but in the next part. The task structure was: mA1-mB1-rB1—mA2-mB2-rB2-rA1—mA3-mB3-rB3-rA2—mB3-rB3-rA3. Load of A lines again did not affect B lines. Crucially, load of A2 and A3 lines did not affect the recall of A1 and A2 lines, respectively. In Experiment 3, in a design similar to Exp1, time constraint on mA and mB increased the interference across concurrent subtasks. Experiment 4 showed that increasing the similarity between subtask A and subtask B of Exp1 may increase the across-subtask interference. We show that WM information of different concurrent subtasks can be maintained separately, perhaps as part of their goal-directed programs. And, encoding to these non-interfering stores, as well as retrieval from them, might depend on attentional and time-based mechanisms.

Keywords: Working Memory, Working Memory Capacity, Concurrent Tasks, Task-Related Programs, Load-Effect, Interference, Hierarchical Control, Attentional Control, Visual Working Memory.

ÖZET

ÇALIŞMA BELLEĞİ KAPASİTESİ: EŞZAMANLI GÖREVLER BİRBİRİNE KET VURMAK ZORUNDA DEĞİL

Gülsüm Özge Şengil

Nörobilim, Yüksek Lisans

Tez Danışmanı: Yardımcı Doçent Dr. Ausaf Ahmed Farooqui

Ekim 2022

Herhangi bir genişletilmiş görev, yürütülmesini hiyerarşik olarak kontrol eden hedefe yönelik programlar tarafından kapsanır. Çalışan bellek (ÇB) kapasitesi ile bu tür bir hiyerarşik kontrol arasındaki ilişkiyi dört deneyde araştırdık. Görev A ve görev B’den oluşan yeni bir genişletilmiş görevde, katılımcılar önce görev A çizgilerinin yönünü ezberlediler (bu olaya mA diyelim), ardından görev B çizgilerini (mB) ezberlediler, sonra bu B çizgilerini (rB), daha sonra da A çizgilerini (rA) hatırladı. Görev yapısı şuydu: mA-mB-rB-rA. Böylece, görev B’nin yürütülmesi sırasında A görevi çizgileri akılda tutuldu. A çizgilerinin artan ÇB yükü yalnızca görev A’daki performansı etkiledi ve görev B’deki performansı etkilemedi. Deney 2’de, Deney 1’in dört denemesi, bir bölümün A çizgilerinin o bölümde değil, bir sonraki bölümde hatırlanması koşuluyla, karmaşık 4 parçalı bir görev haline getirildi. Görev yapısı şuydu: mA1-mB1-rB1—mA2-mB2-rB2-rA1—mA3-mB3-rB3-rA2—mB3-rB3-rA3. A çizgilerinin yükü yine B çizgilerini etkilemedi. En önemlisi, A2 ve A3 çizgilerinin yükü, sırasıyla A1 ve A2 çizgilerinin hatırlanmasını etkilemedi. Deney 3’te, Deney 1’e benzer bir tasarımda, mA ve mB üzerindeki zaman kısıtlaması, eşzamanlı görevler arasındaki ket vurmaya artırdı. Deney 4, Deney 1’in görev A ve B’sinin arasındaki benzerliği artırmanın görevler arası ket vurmaya artırabileceğini gösterdi. Farklı eşzamanlı görevlerin ÇB bilgilerinin, belki de hedefe yönelik programlarının bir parçası olarak ayrı ayrı tutulabileceğini gösteriyoruz. Ve bu karışmayan depolara kodlamanın yanı sıra onlardan geri alma da, dikkat ve zamana dayalı mekanizmalara bağlı olabilir.

Anahtar sözcükler: Çalışma Belleği, Çalışma Belleği Kapasitesi, Eşzamanlı Görevler, Görev Programları, Yük Etkisi, Ket Vurma, Hiyerarşik Kontrol, Dikkate Bağlı Kontrol, Görsel Çalışma Belleği.

Acknowledgement

I would like to thank my advisor Asst. Prof. Ausaf Ahmed Farooqui for his guidance, support, and for giving me the chance of a great exploration during my years here in Bilkent University Neuroscience Graduate Program.

I would like to thank Asst. Prof. Miri Besken, and Asst. Prof. Buse Merve Ürgen for joining the jury and providing warm and valuable feedback.

Bilkent University provided us with great facilities to be able to conduct our research, I would like to extend my thanks to National Magnetic Resonance Imaging Center (UMRAM), and Aysel Sabuncu Brain Research Center, along with the whole family of Neuroscience Graduate Program. This family witnessed my master's degree journey, and I am grateful to get to know many friends who made the experience even more fulfilling. I should also thank TÜBİTAK for supporting me with their bursaries in a 1001 project during my education, and founding this thesis project as part of the grant number 120K924.

Higher Cognitive Functions Lab members created a kind, encouraging, and friendly environment, and they were there whenever I was in need of something. Special thanks to Tamer Gezici who programmed the experiments in this thesis and collected data with me. Our interns and senior thesis students Deniz Özlem Er, Lal Deniz Ersoy, Ece Nur Bilgiç, Ekin Elmalı, and Müge Demirsoy were kind enough to work on the data collection phase of the experiments that consisted of this thesis. Many thanks for your wonderful efforts.

Last, but not the least, I would like to thank my dear family, Gülhan, Melkane, and Buse Can Şengil. I always feel your endless support and unconditional love regardless of the situation I find myself in. I am grateful to have you, always.



To Gülhan, Melkane, and Buse...

Contents

Chapter 1

Introduction

	1
1.1 Working Memory: Definitions and Functions	1
1.2 Theories of Working Memory	3
1.2.1 Memory-systems Models	6
1.2.1.1 The Multiple Component Model	6
1.2.2 State-based Models	8
1.2.2.1 The Embedded Process Model of Working Memory	8
1.2.2.2 Concentric Model of Working Memory	11
1.3 Working Memory Capacity	14
1.3.1 Hypotheses for Why WMC is Limited	15
1.3.1.1 Decay of Information	16
1.3.1.2 Interference	17
1.3.1.3 Resource-Limits	18
1.3.1.3.1 Discrete Allocation of Resources	19
1.3.1.3.2 Continuous Allocation of Resources	21
1.4 Control of Working Memory Information	23
1.4.1 Attention-based Control of Working Memory Information	23

1.4.1.1 Consolidation	24
1.4.1.2 Refreshing	25
1.4.1.3 Removal	25
1.4.2 Hierarchical Control of Working Memory Information	27
1.5 Current Thesis	33

Chapter 2

Experiment 1

	35
2.1 Methods	39
2.1.1 Procedure	39
2.1.2 Participants	40
2.1.3 Experimental Paradigm	41
2.1.4 Data Analysis	46
2.2 Results	47
2.2.1 Subtask B Error	49
2.2.2 Subtask B Recall RT	51
2.2.3 Subtask B Presentation RT	53
2.2.4 Subtask A Error	54
2.2.5 Subtask A Recall RT	56
2.2.6 Subtask A Presentation RT	58

2.3 Discussion	60
----------------	----

Chapter 3

Experiment 2

	66
3.1 Methods	70
3.1.1 Procedure	70
3.1.2 Participants	70
3.1.3 Experimental Paradigm	70
3.1.4. Data Analysis	74
3.2 Results	76
3.2.1 Subtask A WM Load-wise Analyses	76
3.2.1.1 Subtask B Error	78
3.2.1.2 Subtask B Recall RT	79
3.2.1.3 Subtask B Presentation RT	80
3.2.1.4 Subtask A Error	81
3.2.1.4 Subtask A Recall RT	82
3.2.1.6 Subtask A Presentation RT	83
3.2.2 Episode-wise Analyses	85
3.2.2.1 Subtask B Error	87
3.2.2.2 Subtask B Recall RT	88

3.2.2.3 Subtask B Presentation RT	90
3.2.2.4 Subtask A Error	91
3.2.2.5 Subtask A Recall RT	92
3.2.2.6 Subtask A Presentation RT	94
3.3 Discussion	95

Chapter 4

Experiment 3

	104
4.1 Methods	106
4.1.1 Procedure	106
4.1.2 Participants	107
4.1.3 Experimental Paradigm	107
4.1.4 Data Analysis	108
4.2 Results	109
4.2.1 Subtask B Error	109
4.2.2 Subtask B Recall RT	110
4.2.3 Subtask A Error	111
4.2.4 Subtask A Recall RT	112
4.3 Discussion	113

Chapter 5

Experiment 4

	116
5.1 Methods	118
5.1.1 Procedure	118
5.1.2 Participants	118
5.1.3 Experimental Paradigm	119
5.1.4 Data Analysis	120
5.2 Results	121
5.2.1 Subtask B Error	123
5.2.2 Subtask B Recall RT	124
5.2.3 Subtask B Presentation RT	126
5.2.4 Subtask A Error	127
5.2.5 Subtask A Recall RT	129
5.2.6 Subtask A Presentation RT	131
5.3 Discussion	133

Chapter 6

General Discussion

137

6.1 Implications	138
6.2. Further Discussion	141
6.3 Limitations	143
6.4 Future Directions	145

List of Figures

1.1 The schematization of WM theories that are discussed. A. Baddeley & Hitch's Multiple Component Model, with a central executive, episodic buffer, visuospatial sketchpad (VSSP), and the phonological loop. The representations in the phonological loop can be reactivated with articulatory rehearsal. B. Cowan's Embedded Process Model, all the nodes are in LTM, with a subset of them activated (red nodes and orange nodes) in aLTM. A subset of these representations (red nodes) is under the capacity-limited FoA (orange circle). C. Oberauer's Concentric Model, all the nodes are in LTM, with a subset of them being in aLTM (orange nodes). Amongst them, a subset of nodes are in the direct access region (orange circle) related to a common context or mental space (black oval in the bottom, the relation amongst them is shown with orange lines). One item in the direct access region is in the FoA (orange oval).

6

2.1. The Experimental Paradigm of Experiment 1. (A) The order of events and screens in Experiment 1. (B) Different subtasks were indicated by different screen

background colors. Subtask A was indicated by black, and subtask B was indicated by pink screen. (C) Relevant stimuli were colored oriented lines. One sample line per background is shown in the figure, they were taken from the memorization phase of the corresponding subtask. Note that the subtask A line was presented with a circle around it, while the subtask B line was not. **42**

2.2. Sample Screens from Experiment 1. A. Sample mA screen in which 2 A lines are presented. B. Sample mB screen in which 2 B lines are presented. **44**

2.3. Illustration of the results of 2x2 repeated measures ANOVA and planned contrast with Bonferroni correction for Experiment 1, for the output measures of subtask B error (a), subtask B recall RT (b), subtask B presentation RT (c), subtask A error (d), subtask A recall RT (e), and subtask A presentation RT (f). Error bars: 95% CI of the means. The significance stars are only present (or absent) for the pairwise comparisons of interest. **48**

3.1. The Experimental Paradigm of Experiment 2. (A) The order of events and screens in Experiment 2. A session consisted of four episodes. (B) Different subtasks were indicated by different screen background colors. Subtask A was indicated by black, and subtask B was indicated by pink screen. (C) Relevant stimuli were colored oriented lines. One sample line per background is shown in the figure, they were taken from the memorization phase of the corresponding subtask. Note that the subtask A line was presented with a circle around it, while the subtask B line was not. **72**

3.2. Structure of Experiment 2, in comparison with the structure of Experiment 1, 3, and 4. (a) In Experiment 2, the smallest task unit in which the task goals are completed was a session which consisted of four episodes. (b) In Experiments 1, 3

and 4, the smallest task unit in which the task goals are completed was an episode, and each session consisted of four episodes with identical structure. **73**

3.3. Paired Samples T-test Results for Subtask A WM Load-wise Comparisons for Experiment 2. Output measures are: (a) subtask B error (degrees), (b) subtask B recall RT (s), (c) subtask B presentation RT (s), (d) subtask A error (degrees), (e) subtask A recall RT (s), (f) subtask A presentation RT (s). Error bars are 95% CIs of the individual means. **77**

3.4. One-way Repeated Measures ANOVA and Planned Comparison Results for Episode-wise Analyses in Experiment 2. Output measures are: (a) subtask B error (degrees), (b) subtask B recall RT (s), (c) subtask B presentation RT (s), (d) subtask A error (degrees), (e) subtask A recall RT (s), (e) subtask A presentation RT (s). Error bars represent 95% CIs. **86**

4.1. Experimental Paradigm of Experiment 3. A. The order of events and screens in Experiment 3. B. Background colors were different for subtasks. C. Illustrates some sample lines in different backgrounds during the memorization phase. Please, note that the subtask A line was presented with a circle around it, while the subtask B line was not. **108**

4.2. Paired sample T-test results comparing subtask A WM load of 1 and of 2 in Experiment 3. Output measures are: (a) Subtask B Error (degrees), (a) Subtask B Recall RT (s), (c) Subtask A Error (degrees), (d) Subtask A Recall RT (s). **109**

5.1. Experimental paradigm of Experiment 4. A. The order of events and screens in Experiment 4. B. Screen background colors were different for subtask A and subtask B related parts. Subtask A was executed in a pink screen whereas subtask

B was executed in a black screen. C. Illustration of some sample lines in different backgrounds during the memorization phase. Please note that the presentation mode is the same for both of the tasks (neither of them is presented with a circle around them, unlike previous experiments). **119**

5.2. Illustration of the 2x2 repeated measures ANOVA and planned contrast with Bonferroni correction results for Experiment 4, for the output measures of subtask B error (a), subtask B recall RT (b), subtask B presentation RT (c), subtask A error (d), subtask A recall (e), subtask A presentation RT (f). Error bars: 95% CI. The significance stars are only present (or absent) for the pairwise comparisons of interest. **122**

List of Tables

2.1. Descriptives for Subtask B Error (degrees) in Experiment 1	50
2.2. Within Subjects Effects for Subtask B Error (degrees) in Experiment 1	50
2.3. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Error (degrees) in Experiment 1	50
2.4. Descriptives for Subtask B Recall RT (s) in Experiment 1	52

2.5. Within Subjects Effects for Subtask B Recall RT (s) in Experiment 1	52
2.6. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Recall RT (s) in Experiment 1	52
2.7. Descriptives for Subtask B Presentation RT (s) in Experiment 1	53
2.8. Within Subjects Effects for Subtask B Presentation RT (s) in Experiment 1	54
2.9. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Presentation RT (s) in Experiment 1	54
2.10. Descriptives for Subtask A Error (degrees) in Experiment 1	55
2.11. Within Subjects Effects for Subtask A Error (degrees) in Experiment 1	55
2.12. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Error (degrees) in Experiment 1	56
2.13. Descriptives for Subtask A Recall RT (s) in Experiment 1	57
2.14. Within Subjects Effects for Subtask A Recall RT (s) in Experiment 1	57
2.15. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Recall RT (s) in Experiment 1	58
2.16. Descriptives for Subtask A Presentation RT (s) in Experiment 1	59
2.17. Within Subjects Effects for Subtask A Presentation RT (s) in Experiment 1	60
2.18. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Presentation RT (s) in Experiment 1	60
3.1. Descriptives for Subtask B Error (degrees) in Subtask A Load-wise Analysis in Experiment 2	78

3.2. Paired Samples T-test for Subtask B Error (degrees) in Subtask A Load-wise Analysis in Experiment 2	78
3.3. Descriptives for Subtask B Recall RT (s) in Subtask A Load-wise Analysis in Experiment 2	79
3.4. Paired Samples T-test for Subtask B Recall RT (s) in Subtask A Load-wise Analysis in Experiment 2	79
3.5. Descriptives for Subtask B Presentation RT (s) in Subtask A Load-wise Analysis in Experiment 2	80
3.6. Paired Samples T-test for Subtask B Presentation RT (s) in Subtask A Load-wise Analysis in Experiment 2	80
3.7. Descriptives for Subtask A Error (degrees) in Subtask A Load-wise Analysis in Experiment 2	81
3.8. Paired Samples T-test for Subtask A Error (degrees) in Subtask A Load-wise Analysis in Experiment 2	82
3.9. Descriptives for Subtask A Recall RT (s) in Subtask A Load-wise Analysis in Experiment 2	82
3.10. Paired Samples T-test for Subtask A Recall RT (s) in Subtask A Load-wise Analysis in Experiment 2	83
3.11. Descriptives for Subtask A Presentation RT (s) in Subtask A Load-wise Analysis in Experiment 2	84
3.12. Paired Samples T-test for Subtask A Presentation RT (s) in Subtask A Load-wise Analysis in Experiment 2	84

3.13. Descriptives for Subtask B Error (degrees) in Episode-wise Analyses in Experiment 2	87
3.14. Within Subjects Effects for Subtask B Error (degrees) in Episode-wise Analyses in Experiment 2	88
3.15. Planned Comparisons -Episodes - for Subtask B Error (degrees) in Episode-wise Analyses in Experiment 2	88
3.16. Descriptives for Subtask B Recall RT (s) in Episode-wise Analyses in Experiment 2	89
3.17. Within Subjects Effects for Subtask B Recall RT (s) in Episode-wise Analyses in Experiment 2	89
3.18. Planned Comparisons -Episodes - for Subtask B Recall RT (s) in Episode-wise Analyses in Experiment 2	89
3.19. Descriptives for Subtask B Presentation RT (s) in Episode-wise Analyses in Experiment 2	90
3.20. Within Subjects Effects for Subtask B Presentation RT (s) in Episode-wise Analyses in Experiment 2	91
3.21. Planned Comparisons -Episodes - for Subtask B Presentation RT (s) in Episode-wise Analyses in Experiment 2	91
3.22. Descriptives for Subtask A Error (degrees) in Episode-wise Analyses in Experiment 2	92
3.23. Within Subjects Effects for Subtask A Error (degrees) in Episode-wise Analyses in Experiment 2	92

3.24. Planned Comparisons -Episodes - for Subtask A Error (degrees) in Episode-wise Analyses in Experiment 2	92
3.25. Descriptives for Subtask A Recall RT (s) in Episode-wise Analyses in Experiment 2	93
3.26. Within Subjects Effects for Subtask A Recall RT (s) in Episode-wise Analyses in Experiment 2	94
3.27. Planned Comparisons -Episodes - for Subtask A Recall RT (s) in Episode-wise Analyses in Experiment 2	94
3.28. Descriptives for Subtask A Presentation RT (s) in Episode-wise Analyses in Experiment 2	95
3.29. Within Subjects Effects for Subtask A Presentation RT (s) in Episode-wise Analyses in Experiment 2	95
3.30. Planned Comparisons -Episodes - for Subtask A Presentation RT (s) in Episode-wise Analyses in Experiment 2	95
4.1. Descriptives for Subtask B Error (degrees) in Experiment 3	110
4.2. Paired Samples T-test for Subtask B Error (degrees) in Experiment 3	110
4.3. Descriptives for Subtask B Recall RT (s) in Experiment 3	111
4.4. Paired Samples T-test for Subtask B Error (degrees) in Experiment 3	111
4.5. Descriptives for Subtask A Error (degrees) in Experiment 3	112
4.6. Paired Samples T-test for Subtask A Error (degrees) in Experiment 3	112
4.7. Descriptives for Subtask A Recall RT (s) in Experiment 3	113

4.8. Paired Samples T-test for Subtask A Recall RT (s) in Experiment 3	113
5.1. Descriptives for Subtask B Error (degrees) in Experiment 4	123
5.2. Within Subjects Effects for Subtask B Error (degrees) in Experiment 4	124
5.3. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Error (degrees) in Experiment 4	124
5.4. Descriptives for Subtask B Recall RT (s) in Experiment 4	125
5.5. Within Subjects Effects for Subtask B Recall RT (s) in Experiment 4	125
5.6. Post Hoc Comparisons -Subtask A Load * Subtask B Load - for Subtask B Recall RT (s) in Experiment 4	126
5.7. Descriptives for Subtask B Presentation RT (s) in Experiment 4	126
5.8. Within Subjects Effects for Subtask B Presentation RT (s) in Experiment 4	127
5.9. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Presentation RT (s) in Experiment 4	127
5.10. Descriptives for Subtask A Error (degrees) in Experiment 4	128
5.11. Within Subjects Effects for Subtask A Error (degrees) in Experiment 4	129
5.12. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Error (degrees) in Experiment 4	129
5.13. Descriptives for Subtask A Recall RT (s) in Experiment 4	130
5.14. Within Subjects Effects for SubtaskA Recall RT (s) in Experiment 4	130
5.15. Post Hoc Comparisons - Subtask A Load *Subtask B Load - for Subtask A Recall RT (s) in Experiment 4	131

5.16. Descriptives for Subtask A Presentation RT (s) in Experiment 4	132
5.17. Within Subjects Effects for Subtask A Presentation RT (s) in Experiment 4	132
5.18. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Presentation RT (s) in Experiment 4	132



Chapter 1

Introduction

1.1 Working Memory: Definitions and Functions

Working memory (hereafter, "WM") refers to the ability to maintain goal-relevant information from long-term memory or external environment in a highly available state and manipulate this information to flexibly support current cognitive processing [1; 2]. A real-world example given by Adams, Nguyen, and Cowan [3] is like the following: A teacher tells the class that the Earth is the third planet in the solar system and then chooses a student to show the place of the Earth in the solar system map they have on the wall. The student needs to keep the teacher's speech about the order of the Earth in the solar system in mind while processing the question. She needs to start counting from the correct planet, not from the sun, for example, and by monitoring her counting, she should remember to stop at three. Simultaneously, other thoughts such as the anxiety of performing in front of the class might be preoccupying WM. The authors suggest that this complex process can go wrong in many ways leading to WM errors simply because many processes compete for WM capacity, which was said to be limited.

Although giving real-world examples might be practical to convey an idea of what WM means, no single definition would encompass all its different theories and applications [3]. Different researchers with different research agendas widely use the term; hence, there are various definitions. Nonetheless, WM is frequently suggested to possess characteristics such that it holds a 'limited' amount of

information [3]. It plays a uniquely crucial role in cognitive processes [3], and it helps to carry out ecologically relevant tasks in a hierarchically controlled manner [4]. WM information is separate from other memory [3]. WM possesses the ability to shift between a maintenance and transformation mode allowing the representations from long-term memory [hereafter, LTM] or current perception to rapidly bound together with each other and with the representations of the context to form new representations [1].

Any conceptualization of WM assumes some functions to it. To give a set of functions that might apply to many contemporary conceptualizations of WM, we can turn to a recent paper by Postle and Oberauer [1]. There it is pointed out that the function of WM is to hold information available for processing in a controllable manner and not to remember the information simply. They define some other characteristics of its function. The operation of WM requires a medium for temporary binding of representations that are already existing in LTM (chunks) to reorganize them into new representations. These new representations will be representations ‘in working memory.’ Secondly, manipulation of the representations in WM is necessary. It is suggested that such manipulations should at least include a mechanism to selectively access the representations and a mechanism to maintain the rules, goals, or plans to manipulate these selectively accessed representations. Another requirement of WM is said to be selectively maintaining the information most relevant to the current goal. This requirement creates a conflict between the two demands, namely the need for WM representations to be protected from the influence or interference from perception or LTM and the need for these representations to be updated rapidly with the

information coming from perception or LTM. So, WM is suggested to work in a fashion that allows it to switch between these two seemingly conflicting modes to overcome this conflict.

One might ask why WM is a popular subject to study in human cognitive psychology. First, WM is thought to contribute to a wide range of high-level cognitive functions such as cognitive control and planning [1]. Moreover, WM measures correlate with intellectual measures (much better compared to short-term memory [STM] measures), such as fluid intelligence [5]. In addition, WM impairment is a characteristic of many neurological and psychiatric disorders (psychotic disorders such as schizophrenia, schizoaffective disorder, bipolar disorder with psychosis [6], attention deficit hyperactivity disorder (ADHD) [7], Alzheimer's Disease [8], Parkinson's Disease [9], Major Depressive Disorder [10]). Today, working memory is demonstrated and studied widely, with all sensory modalities; amongst them, visual and auditory modalities and a linguistic/verbal approach are used most extensively [1]. It would still be fair to state that both conceptual and empirical issues are lingering around the concept, and an abundance of theoretical approaches is present.

1.2 Theories of Working Memory

The first use of the WM concept was for computers, which arose from the need to store relevant information in problem-solving routines [3]. Computer WM was defined as a separate store to hold possibly multiple concurrent information briefly and temporarily to assist the current tasks [11]. With the emergence of the cognitive revolution, the idea that cognition is an information-processing led to

the understanding that the human mind should have an aspect or mechanism that resembles such structures and mechanisms in computers [1]. Inspired by the era, in the 1960 book of Miller, Galanter, and Pribram, the concept of WM is introduced in the context of cognitive psychology of humans for the first time [3; 4; 1]. Their understanding was that working memory helped to carry out 'plans' by gathering 'data' from outside and within the organism [4]. 'Plans' were the equivalent of computer programs in humans; they were hierarchically organized processes that could control the order of an organism's operations. When any part of a plan was 'executed,' authors suggested, the relevant parts of the plan were consciously accessible and coordinated with some other plans, and for these reasons, they were probably put into either a special place or into a special state for them to be remembered during their execution. Without specifying the mechanism of any sort for such memory, they referred to this quick-access memory that executes plans as 'working memory.' In their conception, several plans or sub-parts of a single plan could be maintained in the working memory at a given time. Furthermore, for a single 'plan,' there would be a hierarchical structure including sub-plans, and sub-sub-plans, with the lowest element controlling the momentary actions [3].

Since the introduction of the term WM into the field of human cognitive psychology, there have been many theories, models, or frameworks to explain the phenomena both theoretically and empirically. It is not possible to directly compare different theories of WM (or of any other concept, for that matter) because although seemingly working on the same concept, different theories empirically study the subject matter based on often different definitions of it [3].

Still, an attempt to divide these models or frameworks into two categories based on how Miller, Galanter, and Pribram [4] conceptualize the operation of WM can be made [12]: There are two possibilities regarding the status of WM information, it is either put into a 'special place' or a 'special state.' If WM contents are 'put into some special place,' it is implied that there is a dedicated mechanism for WM. If, on the other hand, WM contents are 'put into some special state,' this special state is a state or property of the system that explicitly represents the information of interest. That state can transiently change when the representational aspects of the system that supports a cognitive process (such as language or sensory perception) change cognitively or neurally (such as synaptic changes in weights) [1]. The state changes through the operation of this system to the ends and by means that do not make immediate working memory demands (e.g., the system might be simply perceiving) [1]. It is possible to place some influential existing theories or frameworks into either category. This categorization is based on the substantial question of how WM works, more specifically, whether WM has a specific mechanism for its function or emerges from the other mind/brain processes or cooperation of other cognitive and neural systems. Some early theories of WM can be put into the first category since they emphasize distinct memory systems [14]. For this reason, they will be referred to as 'Memory-systems Models.' Some later theories, however, mention a common store, and different memory representations (such as short-term- or long-term memory representations) result from differential activation states within this store [15]. For this reason, they will be referred to as 'Activation-based Models' or 'State-based Models' [15].

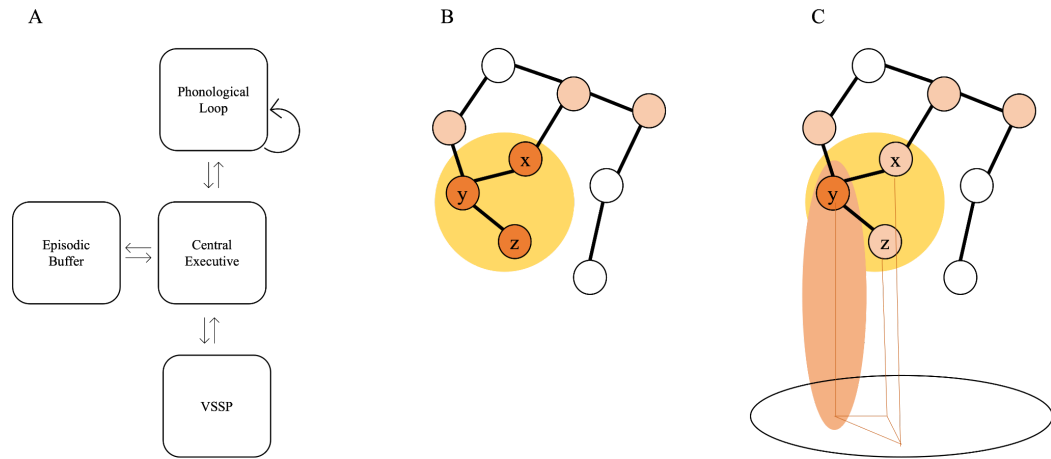


Figure 1.1. The schematization of WM theories that are discussed. A. Baddeley & Hitch's Multiple Component Model, with a central executive, episodic buffer, visuospatial sketchpad [VSSP], and the phonological loop. The representations in the phonological loop can be reactivated with articulatory rehearsal. B. Cowan's Embedded Process Model, all the nodes are in LTM, with a subset of them activated [red nodes and orange nodes] in aLTM. A subset of these representations [red nodes] is under the capacity-limited FoA [orange circle]. C. Oberauer's Concentric Model, all the nodes are in LTM, with a subset of them being in aLTM [orange nodes]. Amongst them, a subset of nodes are in the direct access region [orange circle] related to a common context or mental space [black oval in the bottom, the relation amongst them is shown with orange lines]. One item in the direct access region is in the FoA [orange oval].

1.2.1 Memory-systems Models

1.2.1.1 The Multiple Component Model

The multiple component model is arguably the most influential WM model in the term's history [1]. The model was first proposed by Baddeley and Hitch [16], and since then, it has been occasionally updated. Before this model was proposed, the

influential STM model of Atkinson and Shiffrin assumed a hypothetical single store that temporarily holds information, of which operation is responsible for the performance of many STM paradigms of the day [17]. Not only that, STM was attributed some central role in controlling executive functions or systems in learning and retrieving information [18]. However, Baddeley and Hitch [16] pointed out some empirical evidence to show that STM does not seem to serve a crucial working memory function as it is conceptualized, and they claimed that a common working memory system for reasoning, comprehension, and learning should abandon the single module approach. Abandoning unitary-store short-term memory models of the day, the ‘multiple component model’ emphasized the role of such systems on cognitive operations rather than exclusively on memory [19]. In this model, three domain-specific components to hold different types of information are defined; phonological loop for verbal information, visuospatial sketchpad for visuospatial information, and episodic buffer for episodic long-term memory information [19]. The fourth component is a processor called the central executive; it controls and monitors the ‘slave systems’ (other components) and retrieves information from episodic and semantic LTM [20] (Figure 1.1.A). It is suggested that the operation of the central executive resembles more of a control system over attentional processes.

In a complex span task, a paradigm that is widely used to study WM in Baddeley and Hitch’s book, a list of items is held in mind, then another task (such as a mental arithmetic or reasoning task) is completed, and finally, the first list of items are recalled [16]. Baddeley and Hitch suggested that in such a paradigm, interference between the original items held in mind and the second distractor task

should only happen when the two tasks depend on the same component, which is to say when the two tasks use the same type of information codes stored on the same domain-specific store at the same time [16; 3]. For example, when two verbal or spatial tasks are to be carried out, interference is expected; however, when a verbal task and a spatial task are presented within the same complex span task, interference is supposedly not to occur [3].

It is essential to note the nature of the proposed relationship between the WM and LTM in such a model. When the model was first proposed, it was implied that the central executive had its own multi-modal memory that works beyond the modality-specific operations of the other components [16; 21]. In the following years, this memory of the central executive was first eliminated and then added back to the model under the name of episodic buffer [20]. Such a component (or such an aspect of the central executive that was previously eliminated) was crucial to explain what happens to semantic information which does not seem to fit into any domain-specific component and, most importantly, to give an account for the relationship between the proposed WM system and LTM [21]. Although this model has seen WM and LTM as functionally separate memory systems, cross-modal operations of WM were heavily dependent on LTM, such as retrieving the links between two aspects of an object [21], recombination of LTM information to interpret them in novel ways, finding novel information or solutions from existing LTM information, and encoding the output of WM processes to LTM [22].

1.2.2 State-based Models

1.2.2.1 The Embedded Process Model of Working Memory

What characterizes the information to be held in mind temporarily and makes it distinct from long-term memory has been discussed since early human memory research. Regarding the information transfer from LTM to temporary storage, some early frameworks stated that temporary memory is equal to the information within the LTM but in an elevated state of activation [23]. For example, Hebb's neurologically plausible memory framework suggested that temporary memory is the sum of information from activated cell assemblies [24; as cited in 26]. In his model of memory and attention, Norman also suggested [26] that stimuli can automatically activate the concepts from LTM. Some other frameworks emphasized a crucial theoretical question of whether memory can be activated outside of conscious awareness and suggested that the contents of short-term storage are equal to the contents within conscious awareness [23; 27]. 'Primary memory' of William James [28] was such a concept; it would hold the information for a certain amount of time by making such information conscious. One possible solution to this theoretical problem would be to say that short-term memory and awareness are synonymous and that the level of activation is the sole determinant of awareness.

Although there are many other views followed from the early work that makes unclear comments regarding the distinction between the focus of attention and activation aspects of STM or WM information, in his 'embedded process model,' Cowan argued that such distinction is crucial and WM refers to all the information that is currently accessed or used through an embedded organization of the focus

of attention (FoA), activated long-term memory (aLTM), and LTM [27]. Such information was embedded in a way that memory in the focus of attention was a subset of the aLTM, and aLTM was a subset of the LTM [27] (Figure 1.1.B).

FoA is a capacity-limited state that holds the conscious information to manipulate it, whereas aLTM information would be outside of conscious awareness or attention. Attentional focus is directed to a subset of activated memory mainly by two mechanisms, either automatically due to the bottom-up characteristics of the information [such as a physical change in the stimuli or encountering a stimulus that is of special importance] or by the central executive processes with the influence of rules or instructions. A key aspect of this model is how representations in aLTM and FoA are distinct from each other. FoA is said to be capacity-limited with a discrete limit of 3-4 chunks [29; 1], whereas aLTM has no such item-related capacity limits; its operation is only said to be affected by interference and passage of time [30]. Processing in FoA does not get affected by the information held in aLTM [12]. Moreover, the information in aLTM is assumed to be subject to decay through time and interference, whereas the information in FoA is protected against forgetting [1]. Also, as implied before, representations in aLTM consist of already existing information that is activated, while individual representations in FoA can be bound together to create novel ones [27].

There are no multiple modules in this model; however, the prediction regarding the interference is similar to what the multiple component model would also predict. It is suggested that interference is to occur more when the items that are

held have more features in common, not because they are held in the same store, but because in the activated portion of LTM, concurrently held items with similar features will cause more interference due to their reliance on the same neural apparatus for their features [3].

1.2.2.2 Concentric Model of Working Memory

For some time now, it has been clear that both storage and processing functions are attributed to WM, and how these functions interact with each other remains to be a crucial question. While the simplest explanation for such interaction requires a single resource for both storage and processing [31], evidence pertaining to this interaction suggests that there are situations in which the processing of information does not get affected by holding some other information in mind [32; 33; 34]. Such evidence usually comes from dual-task studies or complex-span WM tasks in which a memory task is interleaved with another task that requires processing so that the dual nature of measuring WM is emphasized; memorization and manipulation [31] (remember that such paradigm was also heavily used in Baddeley & Hitch's book to focus on the distinguished operations of storage components and central executive). Cowan's embedded WM model is, in fact, a tribute to such findings because it suggests that the information that is in hand for immediate access and use (FoA) and the information that is held in the background to be used in further processing (aLTM) are distinguished from each other with the help of selective attention [27; 12]. This conceptualization allows for information in aLTM to be held in the background without interruption from FoA [27; 31; 12]. Therefore, selective attention or FoA seems to be one possible

theoretical resolution to the findings regarding seemingly independent storage and processing in some complex span WM studies.

The idea of FoA or attentional focus on WM seems relatively tidy; however, Oberauer notes different interpretations of FoA in different empirical studies. In Oberauer, Demmrich, Mayr, & Kliegl's study [34], participants needed to perform a complex-span task in which they needed to hold three to six digits in mind while concurrently working on some mental arithmetic problems. The list of digits and the mental arithmetic task were utterly independent in one condition. In the other condition, participants needed to replace some digits in the arithmetic task with the digits they were holding in their minds. The information held in mind during the mental arithmetic task only affected the performance when the withheld list and arithmetic task were not independent of each other. They suggested this kind of finding could be interpreted in two ways based on Cowan's embedded WM theory: during the mental arithmetic task that requires the additional use of information that is held in mind (in aLTM), either the whole list is brought into the attentional focus so that the list takes up some capacity of FoA, or one relevant item is selectively brought up to the FoA from aLTM, meaning that larger the list in mind, more difficult to find the relevant item. Another study that bears evidence for the attentional focus in WM was conducted by Garavan [35]. In this study, participants were asked to keep count of the squares and triangles that were presented on the screen, in serial order, independently. Since the counting was self-paced, the reaction times for switching from one mental counter to the other (i.e., from triangle to square) and keeping the count up within the same mental counter could be calculated (i.e., from triangle to triangle). It was

found that switching between mental counters took significantly more time than updating the same mental counter. This finding is interpreted as that one of the counters was held in FoA while the other was outside of FoA, stored or memorized. Although both of these studies present evidence for FoA, one can realize two different functions attributed to it: Oberauer's [34] first explanation for the finding states that the FoA holds several items [almost 4] simultaneously, whereas the second interpretation of the same study, and Garavan's [35] study implies a narrower attentional focus, that is mainly directed to a single object.

Noting two different interpretations of FoA, Oberauer formed his concentric model of WM [12]. WM is conceptualized as the concentric structure of three different functional representational regions. The first is the aLTM which is to memorize information temporarily to be used later [12; 36]. The second region is the region of direct access, which holds several chunks to be used in the current cognitive processing and is capacity-limited [12; 36]. The third is the FoA to hold one selected chunk at a time to be used in the next cognitive operation [12; 36]. Oberauer claims that Cowan's model's capacity limit of 4 ± 1 chunks reflects a capacity limit of the region of direct access in this current model, which corresponds to Cowan's FoA. In a 2002 paper [12], Oberauer gives object-switching as a piece of further evidence for why there should be a narrower attentional focus; within a list supposed to be held in the region of direct access (or Cowan's FoA) when a specific element is reached the RTs for reaching this object is higher if this object is not the object used in the previous operation compared to the object that was accessed in the previous operation [36].

In the concentric model of WM, the capacity limits of WM do not arise from the limited resource-related issues but rather from the hardship of selectively accessing when WM representations are held concurrently [12]. Oberauer mentions two new mechanisms to influence the capacity limits of WM: overwriting and crosstalk [12]. Overwriting refers to the overwriting of similar features in WM when WM representations are similar [12]. Crosstalk, on the other hand, means that when selectively retrieving one of the items from the region of direct access to the FoA, there is competition amongst the elements that are held in the region of direct access (and not within the elements in the aLTM because the elements in the aLTM are not considered for being selected to the FoA) [12; 36].

To summarize, the concentric model of WM sees WM as an organized set of representations with heightened access to current cognitive processes, and within the contents or representations of WM, there are hierarchically differential accesses of the elements that belong to three different functional regions discussed [12; 36] (Figure 1.1.C).

1.3 Working Memory Capacity

It has been mentioned that WM is conceptualized as a limited capacity and how different theories approach capacity issues are addressed to some extent. However, the working memory capacity (hereafter, WMC) deserves its own structured section, for it is intricate with many theoretical and empirical issues regarding WM operations. WMC measures are significant determinants of individual differences in cognitive and intellectual abilities and cognitive

development from childhood to old age. Having a grasp of why WM appears to be capacity limited, therefore, is crucial to understand the limited cognitive and intellectual abilities, their development over the lifespan, and inter-individual differences.

1.3.1 Hypotheses for Why WMC is Limited

The research on WMC is broad, so it is rather practical to decide on the findings that would be informative for why and how WM is capacity-limited. Oberauer, Farrell, Jarrold, and Lewandowsky [37] identify three distinct phenomena that would directly express the presence of WMC across the field.

1. Memory accuracy decreases with the increased amount of information to be held in memory (i.e., set-size or memory load effect). For example, people find it increasingly more and more challenging to recall visual objects as the number of objects in the set to be remembered increases [38]. This finding directly reflects WMC and WM's being capacity-limited, implying that increasing the amount of information will worsen the performance. At some point, the information will surpass the capacity limits.

2. WM information appears vulnerable to forgetting when the presentation of to-be-remembered items or sets is followed by another attention-consuming *distractor* task, such as mental arithmetic or reading aloud, before their recall. In such complex-span or dual task paradigms, the detrimental effect of the distractor task on memory is suggested to be due to the distractor tasks' placing an additional WM load on an already limited capacity that holds the withheld items.

3. WMC measures point out individual differences in capacity limits, and WMC limits show some generality across testing procedures and tasks in correlational studies. Such correlational data suggest there might be a singular concept of WMC.

The theories that would try to account for what makes WM capacity-limited would try to explain these core findings (amongst others); however, as one can guess, not every hypothesis of WMC can explain the same set of related phenomena with the same precision. Nonetheless, different accounts can be summarized to reflect three broad hypotheses for why WM is capacity-limited: decay, interference, and resource limits.

1.3.1.1 Decay of Information

The first hypothesis regarding why WM is capacity limited is that WM representations quickly decay over time. Such hypotheses that would emphasize the role of decay in WM would also need some mechanisms that would protect WM representations from decay. Subvocal-articulation (rehearsal) and attentional-refreshing are said to be such mechanisms. The main prediction of this hypothesis regarding the WMC would be that one can hold information as much as they can rehearse or refresh before they decay to an unrecoverable state. For visual WM, since subvocal articulation cannot be used to rehearse the material, refreshing is more heavily debated. Refreshing, an attention-based mechanism that helps to control WM information will be investigated in more detail in the upcoming sections.

1.3.1.2 Interference

The interference hypothesis of WMC suggests that capacity limitations arise from the interference between the WM representations [37]. I will briefly mention two forms of interference, although other forms are also discussed in the literature.

The first is ‘interference by confusion’ (similar to Obeauer’s ‘crosstalk’ in the section where I discussed his concentric model). This form of interference occurs because there is competition between the items since multiple of them are retrieval candidates. The more activated representation is retrieved, and some theories argue that the items are kept active during the retention interval. In contrast, others argue that they are not active after the presentation and are reactivated in the retrieval phase with the contextual cues that the task provides [37]. Confusion arises if contextual cues are not distinguishable enough to activate the target representation selectively. Contextual cues might be the context information that different sets of information should be maintained in relation to, or within the same context, the cues that distinguish items from each other such as their order in the list or spatial location. So, the more distinguishable the contextual cues for the individual elements of a set of information or different sets of information, the less confusion occurs. The second form of interference is ‘interference by overwriting’ [39]. Similar features of items in WM are overwritten, causing interference. Such interference would be more serious when the items are more similar than distinct.

1.3.1.3 Resource-Limits

The resource is a construct used in cognitive psychology to informally refer to a set of phenomena showing limited accuracy and efficiency of cognitive operations. When a need to define ‘resource’ as an explanatory construct arises, it can be defined as the quantity that enables cognitive and psychological function or process in such a way that the amount of resource allocated to a function or process is directly related to the efficiency or success of them [37]. The use of the concept of ‘resource’ in the WM domain comes with a set of assumptions; maintaining WM information and carrying out other related cognitive operations require some amount or quantity of resources allocated to them, and the success of maintenance and speed and accuracy of related operations are a function of the amount of resource allocated to them.

In visual WM literature, there seem to be two different understandings of the nature of allocating resources to WM representations: discrete allocation and continuous allocation of resources.

1.3.1.3.1 Discrete Allocation of Resources. One of the earliest investigations of WMC was based on the question, 'how many objects can the mind embrace at once?'. Jevon [as cited in 30] rolled some black beans onto a surface and tried to enumerate the exact number without actually counting the beans. He found out that the estimation of three to four items was unerringly done; however, enumeration (subitizing) yielded increased error with larger numbers of items. This was one of the first attempts to understand the limitations of ongoing cognitive processing, and the limits were investigated in terms of the discrete units.

Attempts to understand the limits of temporal storage during the cognitive revolution had a similar fiber. Miller's famous article 'Magical number seven, plus and minus two' [40] promised the possibility of the presence of a constant regarding the fundamental limits of the human mind. In his studies, the number of items that could be recalled in immediate recall tasks and the amount of information that could be apprehended at a single time appeared to be around seven items [40]. However, he also pointed out that the number of items or the amount of information that can be held at a single moment might depend on how the information is grouped together. So, in any such immediate recall case, the limits of recall might reflect not necessarily the limits for individual items but grouped information or 'chunks' [30]. This second point gives unclear messages regarding the theoretical premises of Miller's magical number seven because it does not necessarily follow from this conception that limitations regarding short-term memory are around seven chunks of information [30]. The proposed common limit of seven items might reflect a combined limitation for individual items and chunked information for a given immediate recall task, such as reciting

a telephone number [30]. In fact, the conception of chunk seems like an oversimplification [30]; the most fundamental reason for that is the fact that chunks may be hierarchically organized into super-chunks, and unless knowing exactly how the information is chunked, using the measures of chunking as the measures of WMC might be problematic. Especially theories based on activation-related states would argue that information in WM should not be conceptualized as isolated and simple chunks; instead, the recall should be thought of as if it is from a highly connected web of differentially accessed information with various connections. Nevertheless, the idea of chunk provides the WMC literature with a unit of measurement, the seven-item capacity limit was a rule of thumb for many years preceding the cognitive revolution, and it is suggested that there are seven 'slots' in the short-term memory [30].

Cowan claimed that when a vast majority of evidence is taken into consideration, at least one of the mechanisms pertaining to WM should bear some capacity limits, perhaps FoA, and he, as well, defined these capacity limits in terms of a certain number of items that can be held at a time [30]. When chunking is allowed and controlled, rather than prevented, Cowan argued that three to five chunks of information are that limit [30].

In visual WM literature, the discrete resource allocation idea is present in various forms of 'slot theories.' According to these theories, resources for short-term or working memory maintenance are made up of a limited number of discrete 'slots,' which are allocated to these temporarily held items or chunks [41]. If WM capacity is for N discrete items or chunks, there can be information only

from these N items in the WM at a given time. If the stimulus set consists of more than N items or chunks, then the information regarding a subset of N items can be maintained only [37]. For instance, in a change detection study by Luck and Vogel, participants saw arrays of colored squares [38]. After a brief delay period, they were asked to detect their visual changes (color). As the number of squares increased, participants' accuracy declined (set-size effect), and from the distribution of accuracies, the authors concluded that participants could hold four items of information. Crucially, they saw that this four-item limit was the same even when they asked participants to detect the changes in two features from the same object (color and orientation) instead of one (color). Their interpretation of the results was that this discrete capacity should be understood in terms of conjunctions of features or integrated objects rather than individual features.

1.3.1.3.2 Continuous Allocation of Resources. Capacity limit ideas, as in the 'slot-theories,' often arise from experimental paradigms that use a categorical or discrete stimulus set [42], such as letter recall or digit span tasks [30], and change detection tasks [38] where the stimulus or the change of stimulus is held constant in a discrete manner. For visual working memory, it has been proposed that there are three or four object slots that hold the items in the WM [38], and entering into these slots operates in an all or none fashion [42]; the item is remembered if only if it entered into one of these slots, and in that case, it is remembered with high precision. According to these item-limit views, it is expected to see an abrupt decline in memory once a fixed number of items is held in the memory [42].

Some recent work gave us compelling evidence to reconsider this classical all-or-none view regarding the WM capacity. Studies show that recall precision declines continuously as the number of items held in WM increases [43; 44]. Also, it is seen that more salient or goal-relevant items are remembered with higher precision, even if it means that the memory for other items held is poorer [45]. A resource view that is more compliant with these findings sees WM still as limited in capacity, but the limited resources are distributed amongst the WM items flexibly. This view emphasizes the memory's precision or quality as a way of understanding WM limits instead of the number of items remembered [42]. Without a limit to the number of stimuli to be remembered, there is a limit on the representational medium to be allocated to those stimuli flexibly [42]. There are two basic premises for these resource models of WM: internal WM representations are noisy, and the ratio of noise increases as the number of items in the WM is increased. The more resources are allocated to an item, the more precise its representation. Researchers studied visual WM with the delayed-estimation technique [43], in which both the stimulus and response space are analog (continuous) instead of discrete. This technique is used to study the memory of motion direction, color, and orientation as visual features. In these tasks, recall variability, which is measured as an indication of WM capacity, is calculated as the standard deviation of the error, and it is shown that recall variability increases continuously and gradually as the number of stimuli in the set increases. This finding aligns with the idea that WM resources are shared between the items. Suppose there is an equal distribution of resources. In that case, a normal distribution can be assumed for error variability, as opposed to the item-limit models, which would expect the recall variability to be fixed after the

N number of items in the memory. Within the proponents of the continuous allocation of resource view, the question of how this continuous resource is distributed amongst the WM items has multiple answers. Van den Berg, Awh, and Ma [46] gathered data from multiple studies, compared different models, and suggested that resource is allocated amongst items and trials variably, not equally. Allocation of WMC to its items and the trials of a task, or the control of WMC, is affected by attentional and non-attentional factors that we will discuss in the next section.

1.4 Control of Working Memory Information

1.4.1 Attention-based Control of Working Memory Information

WM's capacity limits, especially those suggested by the discrete capacity views, are seen as evidence of the bottleneck effect in cognitive processing [47]. Central selective attention is suggested to be the mechanism that allows WM operations to be carried out despite its limitations to keep the goal-directed processing from irrelevant aspects of the environment [47]. Kane, Bleckley, Conway, and Engle [48] suggested that the WM capacity of individuals correlated with their attentional control in complex span tasks. It is also suggested that individuals with lower WMC had problems with ignoring distraction during a task requiring WM storage [49]. This coupling between WM and intellectual ability measures might be due to both WM and intellectual abilities requiring attention control [5]; however, more investigation is needed on this relationship [21].

How well one can perform and reason on a given task and what levels of complexity can be tolerated depends on WM capacity and how the WM capacity

is allocated [50]. In recent years, the role of attention in this allocation of WM resources has been discussed, primarily through the lenses of three attentional processes: consolidation, refreshing, and removal [50].

1.4.1.1 Consolidation

Consolidation is suggested to be related to the direction of attention to a memory trace after it is not perceptually accessible anymore and, therefore, temporary stabilization of that memory trace [51–52]. In other terms, working memory consolidation is the process where the temporary sensory representations are transformed into stable memory traces or representations that can be further processed, specifically manipulated, and retrieved even after some delay period [53]. Although consolidation might appear similar to WM maintenance, one crucial characteristic of consolidation is that it is related to the attentional processing of information right after a new transient sensory trace is formed to stabilize its trace in WM maintenance [53].

The first set of evidence regarding the consolidation process comes from attentional blink literature [53]. Attentional blink refers to the phenomenon that when there are 100 to 400 milliseconds between the presentation of two items to be recalled, the recall of the second item is failed [53]. The most classical finding regarding the need for consolidation processes to stabilize the memory representations into WM is that there is a need for a particular critical retention interval after a to-be-memorized item is presented and before it is recalled [53]. A retention interval (the time immediately after the presentation of items till recall) is found to be necessary for consolidation in studies, especially with simultaneous item presentations [54]. Such retention interval is unique in that an interval

presented after some processing does not benefit the memory performance as much as this initial retention interval [51], which is thought to reflect the time course for the consolidation.

1.4.1.2 Refreshing

Refreshing is suggested to be the cycling of attention over all of the current WM contents, regardless of whether they are newly formed or not, to keep them activated and to prevent them from being forgotten [55; 56]. It is a central attention-based mechanism for WM maintenance [57]. This process enables real-time thinking by keeping the selected encoded or retrieved representations of WM active indefinitely [57]. Refreshing, for instance, in a dual-task paradigm, can happen when no attention-consuming processing is going on and prevent WM representations from fading out [57]. Attentional refreshing is suggested to maintain representations of different sensory modalities and multimodal information [58].

Theoretical and empirical work on refreshing suggests two different types. The first form of refreshing happens outside of awareness and very rapidly [57]. The second type of refreshing can be employed more slowly and deliberately, for instance, by consciously thinking of the content of WM or visualizing it in response to a cue (in lab settings) [59; as cited in 57].

1.4.1.3 Removal

Removal is suggested to be the active elimination of no longer goal-relevant information to free up some space in WM, prevent the interference of it with relevant information, and efficiently use the WMC [60; 61; 62]. This process is an adaptive, goal-directed process [13] that helps keep WM contents updated.

Removal of no longer relevant information from WM did not receive much attention until recently because the old popular view was that WM contents are subject to rapid decay due to time [63]. However, for visual WM, for instance, the rate of decay suggested by the studies that favor the passive decay of visual information from WM is so slow that (e.g., 10% of a loss in the recognition ability of the visual WM contents in 12 secs suggested by Ricker, Spiegel, and Cowan [64]) if indeed decay is responsible from the findings that would mean that visual WM should have a capacity limit that is way more than what is suggested by the various views in the literature. Efficient removal of no longer needed information from WM should rely on more rapid processing [63].

Removal of no longer needed or temporarily goal irrelevant information from WM (or capacity-limited parts of WM) is usually studied with the retro-cue paradigm in which participants first encode a set of stimuli and then see a cue indicating which stimulus or a set of stimuli will be relevant or irrelevant for the next operation. Afterward, participants wait for a cue-target interval (a variable amount of time) to see the probe and do the recognition task. Oberauer [65] found a set-size effect of the items cued as irrelevant on the WM operations of the remaining relevant items. Crucially, this set-size effect started to vanish after 1 second, suggesting that after some time, the items cued as irrelevant are removed from the WM completely. New evidence suggests that exclusion of no-longer goal-relevant information from capacity-limited WM can happen either temporarily or permanently depending on whether the information that is now goal-irrelevant will become goal-relevant in the future [13].

1.4.2 Hierarchical Control of Working Memory Information

Goal-directed behavior consists of temporally extended tasks and actions that usually have an intrinsic but variable hierarchical structure to them. Such behavior is made up of smaller subtask entities, such as individual events or smaller acts that are followed sequentially step by step. Previous accounts also recognize the existence of such entities where the information regarding what to do, how, and when to do it are represented as a single entity. Miller, Galanter, and Pribram's 'plans' were such entities. Plans, once formed, control the execution of lower-level steps. For example, the 'plan' of adding sugar to the coffee may consist of sub-plans, and each of these steps might have its own lower-level sub-plans. The crucial idea is that while these sub-plans are being executed sequentially, 'the plan' selects the sub-plans and their lower-level plans sequentially. For example, the behavior of 'adding sugar to the coffee may consist of sequential subtasks or sub-plans such as 'getting a spoon,' 'scooping the spoon to get some sugar,' and 'putting the sugar in the cup.'

Some computational approaches, as well, make use of similar concepts, especially when it comes to studying behavior in ecologically valid settings outside of the laboratory. Suppose we are to explore any hierarchical behavior or routine. In that case, the interrelated components of the action are cast out into a higher level action, and this single new representation is called 'temporal abstraction' [66; 67]. These higher-level abstractions abstract over temporally extended sequential lower-level acts [67]. Once a temporary abstraction is instantiated, its operation continues until a termination point is reached [67].

Such temporally extended goal-directed behavior that requires sequential execution of a set of subtasks are called 'task episodes' by Farooqui and Manly [68]. The emphasis here is not necessarily on where these task episodes stand in the hierarchy but rather on the argument that such task episodes are construed as one cognitive entity even though they consist of smaller subtasks and processes. However, it is not the physical structure of the task but one's construal of it that determines what constitutes a task episode [68]. When this is the case, how a doer construes a task episode can be manipulated with task-irrelevant cues. For instance, the order of the trials and operations or the placement of small and big breaks can constitute such cues. Manipulating these cues might allow one to investigate the task-episode-related phenomena without contaminating it with the factors pertaining to working memory, long-term memory recall, task rules, or experimental paradigms themselves.

Such hierarchical entities, under different names such as plans or schemas, are studied in cognitive control literature; however, it is essential to note that the control embodied by a hierarchical entity, *or task episode*, differs from the control embodied by executive functions or attention. Executive control or attentional control dwells more on the higher-level processes to monitor information and implements lower-level processes to strategize the processing [as cited in 69]. On the other hand, the control that arises from the relationship between the components of a hierarchical structure can be called 'hierarchical control,' and it will organize, control, and run the hierarchical task entities rather than only representing them [68].

Existing models of hierarchical behavior suggest that the hierarchical structure of the given sequential behavior should be directly reflected in the structure of the cognitive system as a hierarchy of 'plans' or 'schemas' [70]. When this is the case, the reflection of such hierarchy into the processing of WM has also been discussed, and some empirical phenomena are suggested to be evidence for it. In sequential item or motor memorization and recall tasks, the first item tends to have larger reaction times (RTs), and the RT for this first item tends to be even larger for tasks that are with more complex or longer sequences [69]. Moreover, when the sequence consists of more than one type of item, the behavioral switch costs of moving on to processing one type of item from another one seem to be decreased when the switch occurs at the sequence boundaries in a task where the sequence of type of item to be processed is predictable [69]. The same is observed for the benefit of repeating the same type of item. To illustrate the last point, we can refer to a paper by Schneider and Logan [69]. Their participants performed judgments of color [C] or shape [S] in a switching task. When a trial required the processing of a different type of information compared to the previous trial [color judgment after shape judgment or vice versa], the error and RTs tended to be significantly higher compared to the trials where the same item type is used as the previous one (color judgment after color judgment, or shape judgment after shape judgment) (the switching cost). Crucial to the task design, participants saw the trials either in the order of CCSS or CSSC in a predictable way. In the order of CCSS, the sequence boundary will also constitute a switch [CCSS-CCSS], whereas, in the order of CSSC, the sequence boundary will be a repeat of the type of trial. When this was the case, they found that the RT and error rate in the first trial of CCSS was not significantly higher than the first trial of CSSC, even

though it constituted a switch trial. Through the lenses of hierarchical behavior execution, the authors explain the findings as follows. Since hierarchical task-sequence representation requires an instantiation at the beginning, the RT and error rate in the first trial tends to be higher compared to the subsequent trials, and this instantiation took longer for complex or longer sequences. In working memory, there is an interplay between the recall processes and this instantiation of hierarchical representations [such as the sequence of trials that is memorized] that would control the identity and lower-level components that would give rise to such empirical phenomena [69].

Investigation of hierarchical control of WM processes has also been made in the context of hierarchical behavior that is unpredictable. In fact, such a context is a better representation of mundane reality where hierarchical behavior is often unscripted and flexible. Farooqui and Manly investigated the possibility that goal-related control processes are still employed to monitor the steps of the operation and goal directed nature of the processing even when there is a lack of precise anticipation of the task episodes [68]. In a task where the same type of trial is repeated several times, these repeated trials of the same type of items are construed as a single task episode. They replicated the results of Schneider and Logan [69]. The presence of the same empirical phenomena even when the hierarchical task-sequence was unpredictable so that the task-sequence representations were absent suggested that rather than being a sign of the hierarchical task-sequence representation in WM, these empirical phenomena might be a sign of temporally extended hierarchical behavior's execution as one cognitive entity [68].

This execution of temporally extended hierarchical behavior as a single cognitive entity or *task episode* should also mean to conceive of the goal of the task episode [68]. In previous accounts as well, conceiving the goal of an action is deemed critical to the governance of the behaviors that would eventually complete the goal [28;71]; this control is attributed to a goal-directed cognitive entity that would monitor and encompass the execution of the task episodes to its goals [72]. Farooqui and Manly suggested that tasks and goals should not be seen as mere linguistic representations of declarative information about behavior or cognition. Instead, they should be understood as the cognitive entities that control the cognitive processes until their completion arrives, and even other executive control processes are instantiated through the mediation of these goal-directed entities in a given task episode [68]. They called these goal-directed cognitive entities 'task episode-related programs', and these programs will subsume any temporally extended sequence of behavior. This will be a term that I will also employ in this thesis. In Farooqui and Manly's study [68], longer item 1 RT and decreased accuracy for item 1 compared to the following items in a serial task, and item 1 RTs and error rates' being even larger for more complex/longer tasks are suggested to be due to the instantiation of these programs at the beginning of the task episode' taking time, and for more complex/longer tasks, even more, larger task programs are required. And in the same article, why sequence boundaries overwrite switch costs is explained in terms of the hierarchical relationship between the task-related program and corresponding lower level trial related configurations (that would be responsible for the switching cost): the task program will subsume the task episode, and the lower level trial related

configurations will be operating hierarchically under the task program so that the trial type changes happening across the episode boundaries and their switch/repeat cost will be changed based on the changes in higher-level task program.

Some neuroimaging studies also provided evidence for the presence of task-episode-related programs. In a 2012 study by Farooqui, Mitchell, Thompson, and Duncan [73], participants viewed a sequence of single letters and searched for four predetermined target letters sequentially. They were told that the search for the first three targets constituted a subtask, and the search for the fourth target constituted another subtask, even though the target detections for these four letters were exactly identical, the only thing that would make them apart from each other was their sequential order in the task episode. The biasing of the task structure was to make participants conceive the first two target detections as a sub-sub-task completion, the third target detection as a sub-task completion, and the fourth target detection as the main task completion. The fourth target detection elicited the greatest activity, followed by the activity elicited by the third target. Activities elicited by the main task completion were greater than sub-task completion, and the activity elicited by sub-task completion was greater than the sub-sub-task completion. Such an activity pattern reflects how the hierarchical organization of task episodes is conceived and how these conceptions result from the change happening in the corresponding level of the execution of cognitive entities or task programs. Lower-level goal completions changed only the execution status of lower-level task programs leading to more minimal activity; however, higher-level goal completions changed the execution status of main task-related programs leading to maximal activity.

Another conclusion that can be drawn from the mentioned conceptualizations is that the cognitive load of the program established at the beginning of the task episode is highest at the beginning of the episode and will decrease as the task episode-related program progresses or lower-level goals are accomplished. Indeed, in a neuroimaging study that replicated the structure of some experiments from the behavioral studies of which results we discussed here [68], Farooqui and Manly [74] observed that the activation of brain regions that are deactivated by cognitive load [i.e., Default Mode Network] is highest at the beginning of the episode and gradually decreases as the episode continues, approaching the baseline activation level. Thus, they showed that starting and finishing task-related programs is related to the task's construal and that its signs can be found in both behavioral and neuroimaging studies.

1.5 Current Thesis

Visual WMC theories and research suggest many ways we can think of the capacity: a magical number of items or chunks (Cowan's four, Miller's seven) or a medium of resources to be allocated to the WM items flexibly and variably. However, one everyday observation persists that we can maintain more information in WM than these views would predict. Not only that, we deal with multiple purposive tasks at once, maintain and process information for them, and manage to control seemingly hierarchical behavior underlying those tasks flexibly. This flexible control of multiple (hierarchical) tasks is so mundane and seemingly effortless that, for instance, at a breakfast table, behaviors such as adding sugar to the tea, reading the newspaper, and talking about one's plans might happen at the

same time, with the components making up these individual behaviors are intricately and possibly hierarchically re-organized all together. These observations are the motivational sources behind this thesis. The first aim of this thesis is to investigate how WMC behaves in relation to the hierarchical control of WM by goal-directed task entities.

Attentional and executive control of a task episode is always instantiated in a context where a goal-directed entity is subsuming the execution of the behavior [68], and instantiation of these processes is suggested to be intermediated by task episode-related programs. However, the interplay between attentional control of a task episode where there are hierarchical and goal-related elements to subsume the execution of processes has never been investigated. The second aim of this thesis is to give an account or start a discussion of this interplay between attentional and hierarchical control of WMC.

Chapter 2

Experiment 1

The structure of the ongoing hierarchical behavior execution is reflected in the processing of WM information [68; 69]. WM operations showed behavioral signs of the installation and boundaries of hierarchical task entities [68] and neural signs of the beginning and the end of such task episodes [74]. However, prior to this thesis, investigation of the relationship between Working Memory Capacity (WMC) and the execution of hierarchical behavior in WM was somewhat limited.

In Experiment 1, to reveal the interactions between the hierarchical execution of WM operations and WMC, we created a new temporally extended visual WM experimental paradigm in which two WM tasks (i.e. subtask A and subtask B) with very similar demands are hierarchically organized into an extended task episode. A task episode consisted of memorizing subtask A WM information, memorizing subtask B WM information, recall of subtask B WM information previously memorized, and finally, recall of subtask A WM information previously memorized at the beginning of the extended task episode, respectively. In other words, the execution of subtask A is interleaved with the execution of subtask B. The ‘task episode’ in which the higher goals related to the hierarchical execution of the goal-related behavior are completed in this newly introduced paradigm consists of the completion of subtask A and subtask B that are hierarchically organized.

Here are the crucial properties of the newly created extended task episode in Experiment 1:

1. The subtasks that are hierarchically organized into an extended task episode are both WM tasks; they require the maintenance of a set of relevant stimuli, and manipulation of them in accordance with the current task goals.

2. The subtasks that are interleaved have almost identical task demands in terms of the nature of the stimuli to be stored, the operations to be performed, and the nature of the overarching goal of the subtask [memorizing a set of visual information of the same type and retrieving one of the contents when probed]. Therefore, the subtask representations should be very similar in WM, and the task-related programs that are instantiated by the subtasks should be almost identical. The current experimental paradigm is unique because it consists of two subtasks that are identical with similar task demands and are hierarchically organized under an extended task episode.

3. The current design of the extended task episode allows us to independently manipulate loads of both subtasks. This way, we can investigate the WM load effects in the current design. Memory-load or set-size effect is the common finding in the literature that increasing the amount of information will worsen the performance, and it is said to be the main finding in the literature that is suggested to directly reflect that the WM is capacity limited. The implication of memory-load or set-size effect is that the items in the set to be held in WM depends on the same WM resource and since that resource is capacity-limited, increasing the amount of information in a set causes either more depletion of limited resources or decreased amount of resource that is allocated to a given item. And, this in turn, is reflected as worsening recall performance for the items

when the WM information that is held in mind is larger. With the current experimental paradigm that allows us to independently manipulate loads of both subtasks, we can investigate whether the WM load of a given subtask affects the performance in the respective subtask, interferes with the other subtask's load or affects the performance of the other subtask. These investigations can be interpreted in terms of WMC issues that the new paradigm has to offer.

4. In the extended task paradigm that is created, visual stimuli are used. Crucially, both the stimulus and response space are continuous rather than discrete. This aspect of the newly introduced paradigm is in line with the recent research that emphasizes the continuous allocation of WM resources to the WM representations. This will also allow us to observe the set-size or load effects by assessing the precision of memory by referring to continuous measures of accuracy and RT rather than categorical. This approach is found to be more sensitive to the WMC study by the accounts that emphasize the memory's precision or quality as a way of understanding WM limits instead of the number of items remembered [42].

When these points above were taken together, the extended task paradigm that we introduce in Experiment 1 allowed us to investigate the relationship between WMC and control instantiated by the task entities that are hierarchically organized and concurrently executed.

A classical dual-task paradigm is the primary paradigm used extensively in WMC research, for it emphasizes the dual nature of testing WM, namely the maintenance and manipulation [31]. The extended task episode created in

Experiment 1 to reveal the relationship between hierarchical control of WM and WMC shows many departures from a classical dual-task paradigm:

1. Dual-task paradigm, even though it is used to study WMC, has been rarely used to study the mutual effects of the tasks on each other; the emphasis is on the effect of the distractor task on WM tasks, and not vice versa. However, the newly introduced task paradigm would allow us to investigate the mutual effects of both concurrent tasks on each other. In fact, since the subtasks bear remarkable similarity to each other, the mutual effects of the concurrent tasks on each other reflect more of the effect of their hierarchical organization rather than a WMC issue resulting from the unique demands of a given task.

2. In classical dual-task studies, either attentional bottleneck or capacity limits are investigated with the distractor task mostly being an attention task [not a working memory task], so that in a classical dual-task context, the effect observed is rarely the effect of a second WM task on another WM task, let alone the second WM task of the exact same nature.

3. The WM tasks extensively employed in dual-task paradigms share a common characteristic that both the stimulus and response space are either categorical or discrete. This leads to an all-or-none understanding that results in conceptualizing the WMC in terms of the number of items remembered instead of considering the possibility of continuous allocation of WM resources.

All in all, the extended task paradigm that is created for the current thesis consists of two visual WM subtasks that are hierarchically organized so that the information from one WM subtask is withheld while performing another subtask, and then the withheld information is retrieved. The current experimental paradigm

allows us to investigate how WM information belonging to different hierarchically organized subtasks of the exact nature that reside under the same extended task episode is handled in terms of the WMC capacity.

The aim of Experiment 1 was to explore the relationship between the hierarchical control of WM and WMC by using this newly introduced extended task paradigm. More specifically, investigations regarding the effect of withheld subtask A WM load on the concurrently processed subtask B performance, and effect of subtask B WM load on the subtask A performance were the most crucial to this aim.

2.1 Methods

2.1.1 Procedure

The experimental procedure was approved by the Research Ethics Committee of Bilkent University. A convenience sampling method was used, participants were recruited by sending recruitment emails to university's email lists. The criteria for being a participant in this experiment was being 18-65 years old, having normal or corrected to normal vision, and not having a neuropsychological disorder.

Our experiments were conducted online via the platform Pavlovia (www.pavlovia.org). Experiments were coded using the PsychoPy builder [75]. Python code generated via the builder was converted to JavaScript by using the PsychoJS library, experiments could be conducted on a web server hosted by Pavlovia. Experimenters met each participant individually in a prearranged Zoom meeting, during which they were instructed about the procedure and the rules of

the experiment, and given informed consent to be signed digitally. They were let know that their time would be compensated with either course credit or ₺15-25 based on their performance on the task.

Experimenters provided the participants with the experiment link, they would browse to this link and run the experiment locally on their personal computers. After sufficient practice under the observation of the instructor via a screen share, they would be provided the main experiment link, and end the Zoom meeting. Participants did experiment at their pace with break prompts after each session. Data was generated and collected through the Pavlovia server. After, participants would send an email to the experimenter, stating that they are done with the experiment and their choice of compensation. If they preferred to have monetary compensation, the fee was sent to their bank accounts. Finally, they were debriefed about the experiment via emails.

2.1.2 Participants

45 healthy participants (27 females) within an age group of 18-55 ($M=24.5$, $SD=2.579$) were recruited. Participants were given digital copies of the informed consent, and they signed them with their e-signature before participation in the experiment. Participants were reimbursed with course credits or with a payment of ₺15-25 based on their performance. We had two cutoff points in mean error score, participant's mean errors were calculated right after the experiment, and if they wished to receive the monetary compensation, the payment was made based on these scores as ₺15, 20, or 25. Prior to participation, participants were asked about their biological sex and date of birth for descriptive statistics.

2.1.3 Experimental Paradigm

It is mentioned in the introduction that recent evidence suggests that the WMC should be seen as a resource that is flexibly allocated to the items or their features. This type of evidence usually comes from the delayed-estimation studies in which the stimulus and response spaces are continuous rather than discrete. One such study that uses a derivation of delayed estimation technique was conducted by Gorgoraptis, Catalao, Bays, and Husain [45]. In their visual WM task, an array of colored oriented lines was presented in various locations on the screen to memorize, and then after a retention interval, ‘probe’ that matches with the color and location of a line from the memory array was shown, but its orientation was randomly decided. Participants were to use a response dial to change the orientation of the probe to match it to the ‘target’ line from the memory array that had appeared in the same location and in the same color. Crucial aspect of the design is that turning the ‘orientation’ response dial would rotate the probe through all 360 degrees. This aspect of the design allowed them to examine the pattern of responses in a continuous way.

In order to be used as the visual WM task that underlies both subtask A and subtask B, we employed a quite similar design in our experiments, as well, with an added element of hierarchical organization of two identical subtasks. In Experiment 1, we created an extended visual WM task in which the relevant stimuli were colored lines’ orientations or obliqueness. Figure 2.1 illustrates the experimental paradigm. Participants memorized the obliqueness of the simultaneously presented lines, and were probed to recall one of them. In each extended task episode, participants first memorized the orientation of subtask A lines [let’s call this event mA], then memorized subtask B lines (mB), then

recalled these B lines (rB), and finally recalled A lines (rA). The episode structure was: mA-mB-rB-rA (Figure 2.1.A)

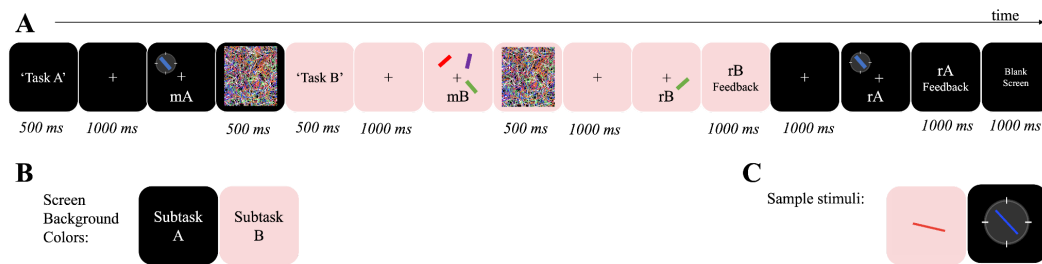


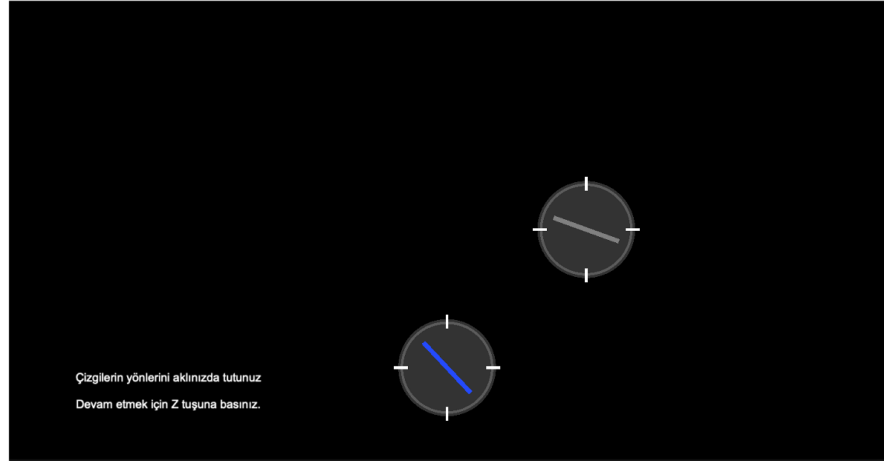
Figure 2.1. The Experimental Paradigm of Experiment 1. (A) The order of events and screens in Experiment 1. (B) Different subtasks were indicated by different screen background colors. Subtask A was indicated by black, and subtask B was indicated by pink screen. (C) Relevant stimuli were colored oriented lines. One sample line per background is shown in the figure, they were taken from the memorization phase of the corresponding subtask. Note that the subtask A line was presented with a circle around it, while the subtask B line was not.

At a finer grain of detail, a given extended episode was like the following: Participants were informed that they can use the background color of the screen as a hint to understand which task they were performing at any moment, the subtask A was indicated by the black background, whereas the subtask B was indicated by the pink/whitish background. First, they saw a black screen with the text ‘Task A’ (for 500 ms) followed by a fixation cross (for 1 s) indicating that they are about to start the new subtask A, then they were presented 1 or 2 lines in different colors and orientations (mA). They were asked to press a button [‘Z’] when they are done with memorization of the orientation of the lines, a visual mask is presented right after the button press. Following the mask, they were presented with a pink screen with a text saying ‘Task B’ (for 500 ms) followed by a fixation cross (for 1

s) cuing that the subtask B is about to be performed. Then, they were presented with 2 or 3 oriented lines on a pink screen for the subtask B memorization phase (mB). After they have indicated the memorization is done by a button press, they again saw the mask (500 ms) and after the fixation (1 s) immediately probed to recall one of the lines that they saw in the previous pink screen of subtask B memorization phase, again on a pink screen (rB). The color and the location of the probe were used to cue the ‘target’ oriented line from the memory array. The orientation of the probe was randomly selected. Participants used arrow keys to rotate the probe to match it to the orientation of the ‘target’, the probe was programmed to rotate 360 degrees with 1 degree intervals. When they were done rotating, they pressed a button (space), and a feedback was given immediately on the same screen, the feedback was given with the target’s original orientation, and their recall deviation from the target’s orientation in degrees (for 1 s). This was concluding the subtask B in the current episode. In the next screen, they saw a black screen with a fixation (for 1 s) indicating that they are about to perform a part of subtask A. After that screen they were probed with a randomly selected subtask A line they saw in the beginning of the episode with a random orientation, and asked to recall its orientation (rA). During any recall, the probe was in the same color and was on the same location on the screen as the target line that is presented, and selected randomly amongst the presented lines. After they indicated the end of the recall, they were shown the orientation of the original target and given the feedback on their recall deviation from the target in degrees (for 1 s). This was concluding the task episode.

Apart from the background color of the screen, we added some other conditional visual features to the lines for them to be more distinctable and indicative of the step of the episode. In mA and rA the lines were presented within a gray circle, in mB the lines are presented freely (Figure 2.1.C), and finally in rB the lines are presented in a purple circle. So that, in a given screen participants knew what was asked and what keys to use to proceed. In addition, the colors of all of the lines presented were different from each other within a given episode, and the different sets of colors are used consistently for subtask A and subtask B lines. Figure 2.2 shows screenshots from the experimental screen for sample mA and mB phases of Experiment 1.

A



B

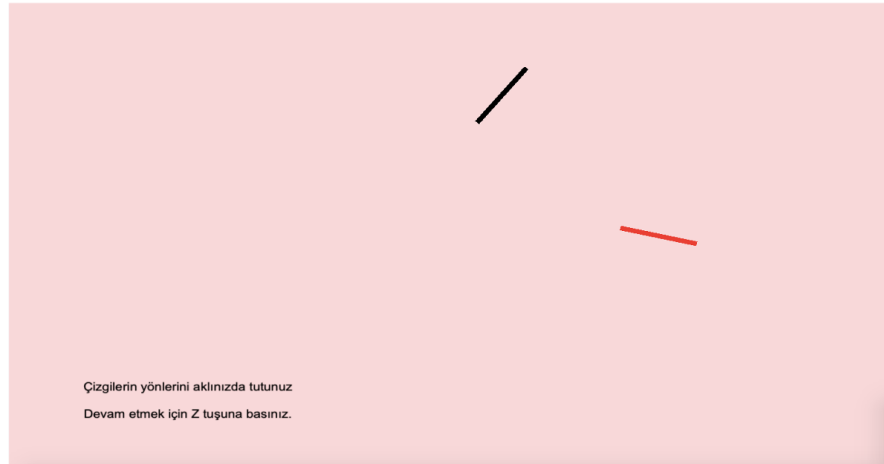


Figure 2.2. Sample Screens from Experiment 1. A. Sample mA screen in which 2 A lines are presented. B. Sample mB screen in which 2 B lines are presented.

As noted, both during the memorization and recall phases, participants were allowed to use as much time as they needed, so the experiment was self-paced. A single run of the experiment consisted of 21 sessions, and each session consisted of 4 episodes. Before starting a new session, and within a session, before starting a new episode, participants saw the notification screen that they are about to start a new episode/session, and in between sessions and episodes, they were allowed to take brief self-paced breaks.

We predetermined four possible locations for the lines to occur with the PsychoPy coordinations of $[0.-3,0.0]$, $[0.3,0.0]$, $[0.0,-0.3]$, $[0.0,0.3]$, these correspond to four dots in the Cartesian system that have equal distance from the center of the screen and are based on two coordinate axes, x and y. For a given memorization phase, locations were chosen randomly from these four in accordance with how many lines to be presented. For the corresponding recall phase, the probe had the same position as its target.

In a given episode, participants were presented with either 1 or 2 lines in mA, and 2 or 3 lines in mB, so that there was 4 different conditions for the number of subtask A and subtask B lines to occur together, respectively, $([1,2], [1,3], [2,2], [2,3])$. The occurrence of such pairs are counterbalanced across episodes within a session. Specifically, 8 different orders that conditions would occur pseudo-randomly in different episodes of a session are created, and these orders

are attributed randomly to participants to ensure within-participants counterbalancing.

Crucial to the design, the number of subtask A lines and subtask B lines are manipulated independently, and subtask A lines were held in mind during the execution of subtask B. To illustrate, if the episode had the [2,3] condition, that meant 2 A lines are held in mind during the execution of B task with 3 lines. Such design allowed us to investigate the effect of withheld WM load (subtask A lines) on a concurrent WM task (subtask B) performance, and vice versa.

2.1.4 Data Analysis

The raw data is prepared for various analyses on MATLAB R2021b, the analyses are run on Jamovi, an R-based computer software [75, 76] and figures illustrating analysis results were created in Prism GraphPad Software [version 9.4.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com] by using the mean and 95% Confidence Interval (CI) information gathered from MATLAB analyses. 95% CIs were calculated by using Cousineau-Morey method for within-participant or repeated measures designs [77; 78]. Descriptive statistics tables, ANOVA tables and pairwise comparisons tables are retrieved from Jamovi [79; 80].

We conducted a 2x2 Repeated Measures Analysis of Variance (ANOVA) to see the effect of subtask A load (categorical variable with 2 levels: 1 line and 2 lines) and subtask B load (categorical variable with 2 levels: 2 line and 3 lines) on all output measures of interest: subtask b error, subtask b recall reaction time, subtask

b presentation time, subtask a error, subtask a recall reaction time, and subtask a presentation time.

As mentioned, we had four pairs of load conditions, and we pre-determined a subset of pairwise comparisons we are interested in for each output measure. After performing planned pairwise comparisons (with Bonferroni correction) following a meaningful ANOVA result, we will be focusing on a subset of comparisons consisting of the following contrasts: ([1,2] vs [1,3]), ([2,2] vs [2,3]), ([1,2] vs [2,2]), and ([1,3] vs [2,3]). In the first two paired comparisons, the subtask a WM load was held constant while the subtask B WM loads differed, and in the last two paired comparisons, the subtask B WM load was held constant while the subtask A WM loads differed. The reason for these comparisons being selected to focus on is that they reflect the effect of manipulating the number of lines that are held in mind for a certain task on the performance of the other task when the latter's WM load was kept constant. The other two pairwise comparisons however, ([1,2] vs [2,3], and [1,3] vs [2,2]) do not provide direct information regarding the effect of WM load of a certain subtask on the other's performance.

2.2 Results

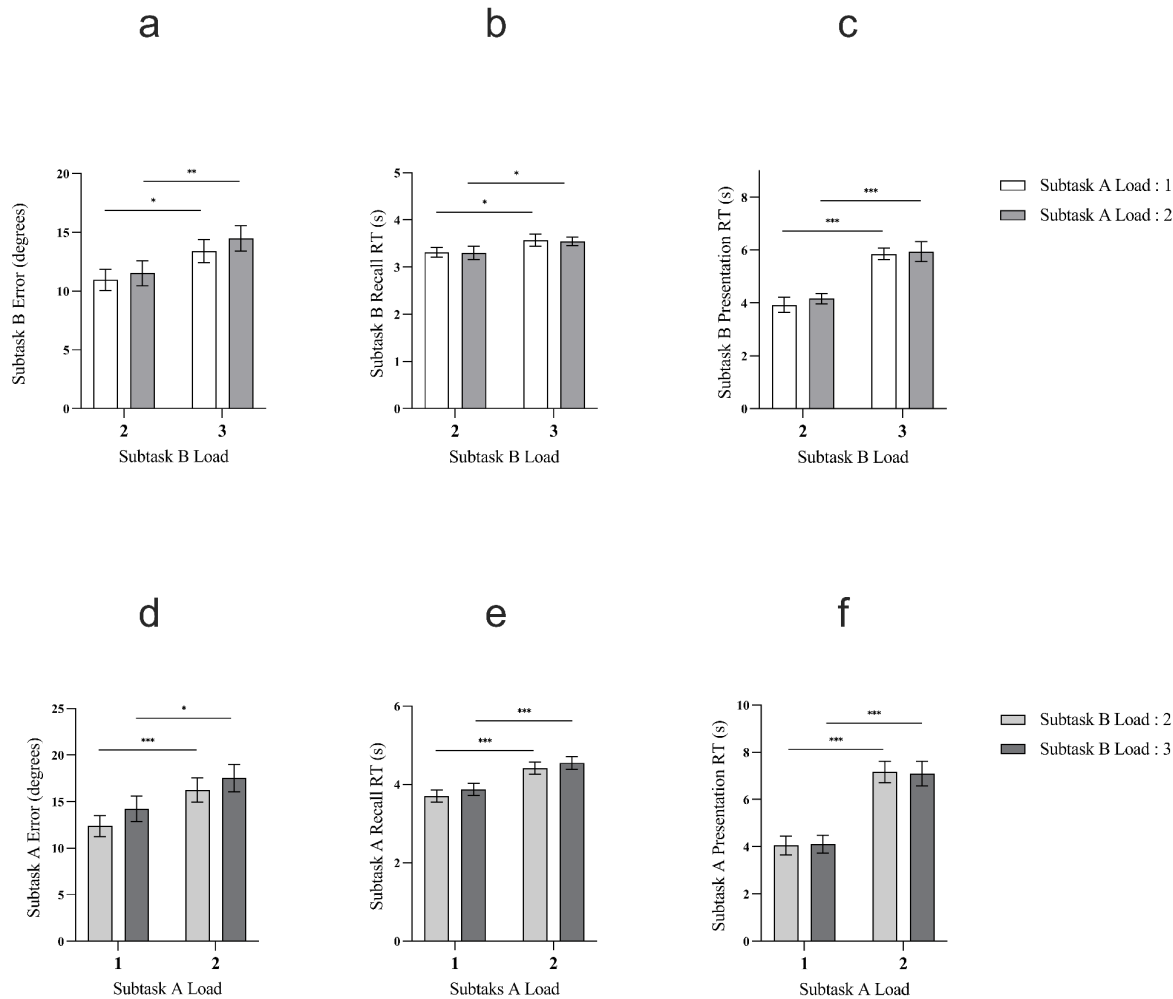


Figure 2.3. Illustration of the results of 2x2 repeated measures ANOVA and planned contrast with Bonferroni correction for Experiment 1, for the output measures of subtask B error (a), subtask B recall RT (b), subtask B presentation RT (c), subtask A error (d), subtask A recall RT [e], and subtask A presentation RT (f). Error bars: 95% CI of the means. The significance stars are only present (or absent) for the pairwise comparisons of interest.

2.2.1 Subtask B Error

Subtask B error was calculated as the mean deviation of the recall in rB from the target shown in respective mB, in degrees. Table 2.1 summarizes the descriptive statistics for subtask B error. The repeated measures ANOVA revealed that (see Table 2.2 for the repeated measures ANOVA summary table) there was only an effect of WM load of subtask B on subtask B error (Subtask B WM Load: $F(1,44) = 35.63, p < 0.001, \eta^2_p = 0.047$), and there was no significant effect of withheld WM load on subtask B error (Subtask A WM Load: $F(1,44) = 2.18, p = 0.15, \eta^2_p = 0.447$). The interaction effect of subtask A load and subtask B load was not significant either (Interaction: Subtask A Load and Subtask B Load: $F(1,44)=0.29, p=0.59, \eta^2_p = 0.007$).

Pairwise comparisons with Bonferroni correction (Table 2.3) revealed that the groups that only differ in their subtask B load, namely, [1,2] vs [1,3] (Mean Difference = -2.45, SE = 0.58, $t(44) = -4.21, p = 0.017$), and [2,2] vs [2,3] (Mean Difference = -2.97, SE = 0.74, $t(44) = -4.04, p = 0.00$) were significantly different. Crucially, the comparisons of the groups that only differed in their subtask A loads, namely [1,3] vs [2,3] (Mean Difference: -1.09, SE: 0.76, $t(44) = -1.43, p=0.97$) and [1,2] vs [2,2] (Mean Difference = -0.57, SE = 0.72, $t(44) = -.80, p = 1.00$) were significantly different. Figure 2.4.a illustrates these results. The results revealed that increasing the number of concurrently executed lines (the subtask B load) affected subtask B error, and crucially, increasing the load of subtasks A did not affect the error on the concurrently executed subtask B WM task.

Table 2.1. Descriptives for Subtask B Error [degrees] in Experiment 1

	1_2	1_3	2_2	2_3
N	45	45	45	45
Missing	0	0	0	0
Mean	11.0	13.4	11.5	14.5
Std. error mean	0.804	0.911	0.932	1.11
95% CI mean lower bound	9.38	11.6	9.69	12.3
95% CI mean upper bound	12.5	15.2	13.3	16.7
Median	8.86	11.7	8.48	12.3
Mode	8.57 ^a	7.33 ^a	6.38 ^a	4.38 ^a
Standard deviation	5.39	6.11	6.25	7.44
Minimum	4.10	4.81	3.57	4.38
Maximum	26.1	27.5	29.2	39.5
Skewness	1.18	0.954	1.20	1.61
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.2. Within Subjects Effects for Subtask B Error [degrees] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A WM Load	30.93	1	30.93	2.180	0.147	0.004	0.004	0.047
Residual	624.37	44	14.19					
Subtask B WM Load	329.34	1	329.34	35.626	< .001	0.044	0.044	0.447
Residual	406.75	44	9.24					
Subtask A WM Load * Subtask B WM Load	3.04	1	3.04	0.289	0.593	0.000	0.000	0.007
Residual	461.91	44	10.50					

Note. Type 3 Sums of Squares

Table 2.3. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Error [degrees] in Experiment 1

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A WM Load	Subtask B WM Load		Subtask A WM Load	Subtask B WM Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-2.446	0.581	44.0	-4.208	< .001
		-	2 [high]	2 [low]	-0.569	0.716	44.0	-0.795	1.000
		-	2 [high]	3 [high]	-3.534	0.695	44.0	-5.089	< .001
	3 [high]	-	2 [high]	2 [low]	1.876	0.748	44.0	2.509	0.095
		-	2 [high]	3 [high]	-1.089	0.764	44.0	-1.425	0.968
2 [high]	2 [low]	-	2 [high]	3 [high]	-2.965	0.735	44.0	-4.036	0.001

2.2.2 Subtask B Recall RT

The subtask B recall reaction time was calculated as the time from the probe is shown till the button response is made in rB, in seconds. Table 2.4 summarizes the descriptive statistics for subtask B error. 2x2 Repeated Measures ANOVA (Table 2.5) revealed that, similar to the subtask B error, there is an effect of WM load of subtask B on subtask B recall reaction time (Subtask B WM Load: $F(1,44) = 15.35$, $p < 0.001$, $\eta^2_p = 0.259$), and there is no significant effect of subtask A WM load on subtask B recall reaction time (Subtask A WM Load: $F(1,44) = 0.10$, $p = 0.75$, $\eta^2_p = 0.002$). Interaction effect between concurrent and withheld WM loads was not significant, as well (Interaction: Subtask A WM Load x Subtask B WM Load: $F(1,44) = 0.34$, $p = 0.86$, $\eta^2_p = 0.001$).

Pairwise comparisons with Bonferroni correction (Table 2.6) revealed that the groups showed a significant difference when the subtask A load is kept constant, for comparisons [1,2] vs [1,3] (Mean Difference = -0.26, SE = 0.08, $t(44) = -3.18$, $p = 0.02$), and [2,2] vs [2,3] (Mean Difference = -0.24, SE = 0.08, $t(44) = -2.91$, $p = 0.03$). When subtask B WM load was kept constant across the pairwise comparisons, however, compared groups did not show a significant difference for the comparisons [1,3] vs [2,3] (Mean Difference = 0.028, SE = 0.07, $t(44) = 0.38$, $p \cong 1.00$), and [1,2] vs [2,2] (Mean Difference: 0.01, SE: 0.08, $t(44) = 0.10$, $p \cong 1.00$). Figure 2.4.b illustrates these results. Similar to the subtask B error, results showed that changing the number of concurrently executed lines affected subtask B recall RT, and crucially, changing the number of withheld lines did not affect the recall RT on the concurrently executed WM task.

Table 2.4. Descriptives for Subtask B Recall RT [s] in Experiment 1

	1 2	1 3	2 2	2 3
N	45	45	45	45
Missing	0	0	0	0
Mean	3.31	3.57	3.30	3.54
Std. error mean	0.224	0.254	0.216	0.240
95% CI mean lower bound	2.87	3.07	2.88	3.07
95% CI mean upper bound	3.75	4.07	3.73	4.02
Median	2.93	3.17	2.75	3.07
Mode	1.37 ^a	1.30 ^a	1.37 ^a	1.52 ^a
Standard deviation	1.50	1.70	1.45	1.61
Minimum	1.37	1.30	1.37	1.52
Maximum	8.46	9.44	7.40	8.61
Skewness	1.80	1.60	1.26	1.60
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.5. Within Subjects Effects for Subtask B Recall RT [s] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A WM Load	0.01465	1	0.01465	0.1037	0.749	0.000	0.000	0.002
Residual	6.21582	44	0.14127					
Subtask B WM Load	2.85588	1	2.85588	15.3509	<.001	0.007	0.007	0.259
Residual	8.18576	44	0.18604					
Subtask A WM Load * Subtask B WM Load	0.00409	1	0.00409	0.0335	0.856	0.000	0.000	0.001
Residual	5.37298	44	0.12211					

Note. Type 3 Sums of Squares

Table 2.6. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Recall RT [s] in Experiment 1

Comparison					Mean Difference	SE	df	t	Pbonferroni
Subtask A WM Load	Subtask B WM Load		Subtask A WM Load	Subtask B WM Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-0.26145	0.0821	44.0	-3.183	0.016
		-	2 [high]	2 [low]	0.00851	0.0807	44.0	0.106	1.000
		-	2 [high]	3 [high]	-0.23388	0.0665	44.0	-3.518	0.006
	3 [high]	-	2 [high]	2 [low]	0.26996	0.1006	44.0	2.682	0.062
		-	2 [high]	3 [high]	0.02757	0.0721	44.0	0.382	1.000
2 [high]	2 [low]	-	2 [high]	3 [high]	-0.24239	0.0834	44.0	-2.908	0.034

2.2.3 Subtask B Presentation RT

The subtask B presentation reaction time was calculated as the mean amount of time participants took to memorize the lines shown in the screen mB, [the time between the mB started and participant made the button response to move onto the next screen], in seconds. Table 2.7 summarizes the descriptive statistics for subtask B presentation RT. 2x2 Repeated Measures ANOVA (Table 2.8) revealed that only the main effect of subtask B WM load was present (Subtask B WM Load: $F(1,44) = 113.38, p < 0.001, \eta^2_p = 0.72$) (Subtask A WM Load: $F(1,44) = 1.65, p = 0.21, \eta^2_p = 0.04$) (Interaction: Subtask A WM Load x Subtask B WM Load: $F(1,44) = 0.53, p = 0.47, \eta^2_p = 0.012$).

Pairwise comparisons with Bonferroni correction (Table 2.9) revealed that when withheld WM load is kept constant between the groups, the mean differences were significant ([1,2] vs [1,3]: Mean Difference: -1.92, SE: 0.017, $p < 0.001$) ([2,2] vs [2,3]: Mean Difference: -1.77, SE: 0.23, $p < 0.001$). When concurrent WM load is kept constant at 2 lines, the comparison of [1,2] vs [2,2] yielded a significant difference (Mean Difference: -0.24, SE: 0.11, $p = 0.04$), however, the comparison of [1,3] vs [2,3] was not significant (Mean Difference: -0.09, SE: 0.20, $p = 0.66$). Figure 2.4.c illustrates these results.

Table 2.7. Descriptives for Subtask B Presentation RT [s] in Experiment 1

	1 2	1 3	2 2	2 3
N	45	45	45	45
Missing	0	0	0	0
Mean	3.93	5.85	4.17	5.94
Std. error mean	0.368	0.486	0.399	0.569
95% CI mean lower bound	3.21	4.90	3.38	4.83
95% CI mean upper bound	4.65	6.81	4.95	7.06
Median	3.11	5.18	3.49	5.25
Mode	0.942 ^a	1.13 ^a	1.03 ^a	1.19 ^a
Standard deviation	2.47	3.26	2.68	3.82
Minimum	0.942	1.13	1.03	1.19

Table 2.7. Descriptives for Subtask B Presentation RT [s] in Experiment 1

	1 2	1 3	2 2	2 3
Maximum	12.6	15.6	12.5	17.0
Skewness	1.31	0.797	1.19	1.13
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.8. Within Subjects Effects for Subtask B Presentation RT [s] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A WM Load	1.201	1	1.201	1.654	0.205	0.001	0.001	0.036
Residual	31.948	44	0.726					
Subtask B WM Load	153.950	1	153.950	113.384	< .001	0.083	0.083	0.720
Residual	59.742	44	1.358					
Subtask A WM Load *	0.254	1	0.254	0.533	0.469	0.000	0.000	0.012
Subtask B WM Load								
Residual	20.953	44	0.476					

Note. Type 3 Sums of Squares

Table 2.9. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Presentation RT [s] in Experiment 1

Comparison					SE	df	t	P _{bonferroni}
Subtask A WM Load	Subtask B WM Load	Subtask A WM Load	Subtask B WM Load	Mean Difference				
1 [low]	2 [low]	- 1 [low]	3 [high]	-1.9248	0.174	44.0	-11.060	< .001
		- 2 [high]	2 [low]	-0.2385	0.112	44.0	-2.124	0.236
		- 2 [high]	3 [high]	-2.0130	0.268	44.0	-7.515	< .001
	3 [high]	- 2 [high]	2 [low]	1.6862	0.144	44.0	11.676	< .001
		- 2 [high]	3 [high]	-0.0883	0.202	44.0	-0.437	1.000
2 [high]	2 [low]	- 2 [high]	3 [high]	-1.7745	0.226	44.0	-7.840	< .001

2.2.4 Subtask A Error

The subtask A error was calculated as the mean deviation of the recall in rA from the target shown in respective mA, in degrees. Table 2.10 summarizes the descriptive statistics for subtask A error. The repeated measures ANOVA revealed that (Table 2.11) both of the main effects were present (Subtask A WM Load: $F(1,44)=21.33$, $p<0.001$, $\eta^2_p = 0.33$), (Subtask B WM Load: $F(1,44) = 7.62$, $p =$

0.008, $\eta_p^2 = 0.15$), but interaction effect (Interaction: Subtask A W Load x Subtask B WM Load: $F(1,44) = 0.21, p = 0.65 \eta_p^2 = 0.005$) was not significant.

Pairwise comparisons with Bonferroni correction (Table 2.12) revealed that the groups showed a significant difference when the subtask B load was kept constant, and subtask A load differed between groups, for comparisons [1,2] vs [2,2] (Mean Difference = -3.87, SE = 0.87, $t(44) = -4.44, p < 0.001$), and [1,3] vs [2,3] (Mean Difference: -3.31, SE: 1.08, $t(44) = -3.06, p = 0.02$). The comparisons of [1,2] vs [1,3] (Mean Difference = -1.85, SE = 0.74, $t(44) = -2.48, p = 0.10$), and [2,2] vs [2,3] (Mean Difference = -1.29, SE = 0.90, $t(44) = -1.44, p = 0.16$), i.e. the comparisons that differed only in subtask B load, did not significantly differ from each other. Figure 2.4.d illustrates these results.

Table 2.10. Descriptives for Subtask A Error [degrees] in Experiment 1

	1 2	1 3	2 2	2 3
N	45	45	45	45
Missing	0	0	0	0
Mean	12.4	14.2	16.2	17.5
Std. error mean	1.44	1.70	1.37	1.60
95% CI mean lower bound	9.55	10.9	13.6	14.4
95% CI mean upper bound	15.2	17.6	18.9	20.7
Median	9.90	11.0	16.4	15.7
Mode	7.62 ^a	5.71 ^a	28.8	13.8 ^a
Standard deviation	9.68	11.4	9.17	10.7
Minimum	2.05	2.43	2.24	1.71
Maximum	47.4	54.7	51.2	49.6
Skewness	2.17	2.18	1.30	1.06
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.11. Within Subjects Effects for Subtask A Error [degrees] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η_p^2
Subtask A WM Load	580.12	1	580.12	21.328	< .001	0.030	0.030	0.326
Residual	1196.80	44	27.20					
Subtask B WM Load	111.12	1	111.12	7.626	0.008	0.006	0.006	0.148

Table 2.11. Within Subjects Effects for Subtask A Error [degrees] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Residual	641.17	44	14.57					
Subtask A WM Load * Subtask B WM Load	3.46	1	3.46	0.214	0.646	0.000	0.000	0.005
Residual	712.60	44	16.20					

Note. Type 3 Sums of Squares

Table 2.12. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Error [degrees] in Experiment 1

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A WM Load	Subtask B WM Load		Subtask A WM Load	Subtask B WM Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-1.85	0.745	44.0	-2.48	0.102
		-	2 [high]	2 [low]	-3.87	0.871	44.0	-4.44	< .001
		-	2 [high]	3 [high]	-5.16	0.946	44.0	-5.46	< .001
	3 [high]	-	2 [high]	2 [low]	-2.02	0.981	44.0	-2.06	0.273
		-	2 [high]	3 [high]	-3.31	1.082	44.0	-3.06	0.022
2 [high]	2 [low]	-	2 [high]	3 [high]	-1.29	0.901	44.0	-1.44	0.949

2.2.5 Subtask A Recall RT

The subtask A recall reaction time was calculated as the time from the probe is shown and button response is made in the screen rA, in seconds. Table 2.13 summarizes the descriptive statistics for subtask A recall RT. 2x2 Repeated Measures ANOVA (Table 2.14) revealed that, similar to subtask A error, both of the main effects were present (Subtask B WM Load: $F(1,44)=8.33$, $p=0.006$, $\eta^2_p = 0.16$) (Subtask A WM Load: $F(1,44)= 44.79$, $p<0.001$, $\eta^2_p = 0.50$), and interaction effect was not significant (Interaction: Subtask A WM Load x Subtask B WM Load: $F(1,44)=0.06$, $p=0.80$, $\eta^2_p = 0.001$).

Pairwise comparisons with Bonferroni correction (Table 2.15] revealed that when the concurrent WM load was kept constant, the different withheld WM load groups showed significant difference in their means ([1,2] vs [2,2]: Mean Difference = -0.71, SE = 0.12, $t(44) = -5.82$, $p<0.001$] ([1,3] vs [2,3]: Mean

Difference = -0.67, SE = 0.13, $t(44) = -5.29$ $p < 0.001$). Crucially, when withheld WM load is kept constant in the pairwise comparisons, none of the comparisons were meaningful ([1,2] vs [1,3]: Mean Difference = -0.17, SE = 0.09, $t(44) = -1.99$, $p = 0.32$) ([2,2] vs [2,3]: Mean Difference = -0.14, SE = 0.09, $t(44) = -1.53$, $p = 0.80$). Figure 2.4.e illustrates these results. While there was a main effect of the subtask B WM load on subtask A recall RT, the pairwise comparisons of interests showed that the means of groups that differ only in subtask B WM load were not meaningfully different from each other. It is possible to find a non-significant pairwise comparison result in a data set where the F test showed a main effect, that is because these two tests answer considerably different questions, and more crucially, the specific pairwise comparisons that we are interested in are not the ones that drive the main effect of subtask B WM load.

Table 2.13. Descriptives for Subtask A Recall RT [s] in Experiment 1

	1 2	1 3	2 2	2 3
N	45	45	45	45
Missing	0	0	0	0
Mean	3.71	3.88	4.41	4.55
Std. error mean	0.196	0.190	0.249	0.271
95% CI mean lower bound	3.32	3.51	3.93	4.02
95% CI mean upper bound	4.09	4.25	4.90	5.08
Median	3.38	3.62	4.02	4.13
Mode	1.86 ^a	1.76 ^a	1.80 ^a	1.33 ^a
Standard deviation	1.31	1.28	1.67	1.81
Minimum	1.86	1.76	1.80	1.33
Maximum	7.86	7.12	9.76	9.39
Skewness	1.14	0.601	1.23	1.01
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.14. Within Subjects Effects for Subtask A Recall RT [s] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A WM Load	21.3396	1	21.3396	44.7906	<.001	0.049	0.049	0.504
Residual	20.9630	44	0.4764					

Table 2.14. Within Subjects Effects for Subtask A Recall RT [s] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask B WM Load	1.0604	1	1.0604	8.3327	0.006	0.003	0.002	0.159
Residual	5.5995	44	0.1273					
Subtask A WM Load * Subtask B WM Load	0.0134	1	0.0134	0.0621	0.804	0.000	0.000	0.001
Residual	9.5161	44	0.2163					

Note. Type 3 Sums of Squares

Table 2.15. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Recall RT [s] in Experiment 1

Comparison					Mean Difference	SE	df	t	p _{bonferroni}
Subtask A WM Load	Subtask B WM Load		Subtask A WM Load	Subtask B WM Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-0.171	0.0857	44.0	-1.99	0.315
		-	2 [high]	2 [low]	-0.706	0.1212	44.0	-5.82	<.001
		-	2 [high]	3 [high]	-0.842	0.1189	44.0	-7.08	<.001
	3 [high]	-	2 [high]	2 [low]	-0.535	0.1127	44.0	-4.75	<.001
		-	2 [high]	3 [high]	-0.671	0.1269	44.0	-5.29	<.001
2 [high]	2 [low]	-	2 [high]	3 [high]	-0.136	0.0890	44.0	-1.53	0.798

2.2.6 Subtask A Presentation RT

The subtask A presentation reaction time was calculated as the mean amount of time participants took to memorize the lines shown in the screen mA, (the time between the mA started and participant made the button response to move onto the next screen), in seconds. In the experiment, because of the very nature of the structure of an episode, mA happens before mB, therefore it is only meaningful to not to expect an effect of subtask B WM load on subtask A presentation RT. However, for the sake of completion, I will still give you the whole picture. Table 2.16 summarizes the descriptive statistics for subtask A presentation RT. 2x2 Repeated Measures ANOVA (Table 2.17) revealed that only the main effect of subtask A WM load was present (Subtask A WM Load: $F(1,44) = 118.07$, $p < 0.001$, $\eta^2_p = 0.73$) (Subtask B WM Load: $F(1,44) = 0.003$, $p = 0.96$, $\eta^2_p < 0.001$)

(Interaction: Subtask A WM Load x Subtask B WM Load: $F(1,44) = 0.23$, $p = 0.64$, $\eta^2_p = 0.005$).

Pairwise comparisons with Bonferroni correction (Table 2.18) revealed that when subtask B WM load is kept constant between the groups the mean differences between groups were significant ([1,2] vs [2,2]: Mean Difference = -3.11, SE = 0.29, $t(44) = -10.58$, $p < 0.001$) ([1,3] vs [2,3]: Mean Difference = -2.98, SE = 0.33, $t(44) = -9.09$, $p < 0.001$). When subtask A WM load is kept constant, on the other hand, the comparisons of [1,2] vs [1,3] (Mean Difference = 0.053, SE = 0.18, $t(44) = -0.29$, $p \approx 1.00$) and [2,2] vs [2,3] (Mean Difference = 0.08, SE = 0.31, $t(44) = 0.24$, $p \approx 1.00$) were not significant, as logically expected. Figure 2.4.f illustrates these results. As one can expect, there is a sole effect of subtask A on subtask A presentation RT considering mA happens before mB. This result can be treated as a control since the subtask A manipulation was effective to create two different WM load conditions.

Table 2.16. Descriptives for Subtask A Presentation RT [s] in Experiment 1

	1_2	1_3	2_2	2_3
N	45	45	45	45
Missing	0	0	0	0
Mean	4.06	4.11	7.17	7.10
Std. error mean	0.346	0.326	0.567	0.574
95% CI mean lower bound	3.38	3.47	6.06	5.97
95% CI mean upper bound	4.73	4.75	8.28	8.22
Median	3.47	3.88	5.94	6.47
Mode	1.30 ^a	1.12 ^a	2.51 ^a	2.16 ^a
Standard deviation	2.32	2.19	3.80	3.85
Minimum	1.30	1.12	2.51	2.16
Maximum	13.1	9.45	18.4	19.3
Skewness	1.65	0.895	1.03	1.08
Std. error skewness	0.354	0.354	0.354	0.354

^a More than one mode exists, only the first is reported

Table 2.17. Within Subjects Effects for Subtask A Presentation RT [s] in Experiment 1

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A WM Load	418.58900	1	418.58900	118.07081	< .001	0.194	0.194	0.729
Residual	155.99042	44	3.54524					
Subtask B WM Load	0.00593	1	0.00593	0.00288	0.957	0.000	0.000	0.000
Residual	90.61928	44	2.05953					
Subtask A WM Load * Subtask B WM Load	0.18695	1	0.18695	0.22515	0.637	0.000	0.000	0.005
Residual	36.53421	44	0.83032					

Note. Type 3 Sums of Squares

Table 2.18. Post Hoc Comparisons -Subtask A Load * Subtask B Load - for Subtask A Presentation RT [s] in Experiment 1

Comparison					Mean Difference	SE	df	t	p _{bonferroni}
Subtask A WM Load	Subtask B WM Load	Subtask A WM Load	Subtask B WM Load						
1 [low]	2 [low]	-	1 [low]	3 [high]	-0.0530	0.181	44.0	-0.293	1.000
		-	2 [high]	2 [low]	-3.1144	0.294	44.0	-10.583	< .001
		-	2 [high]	3 [high]	-3.0384	0.375	44.0	-8.094	< .001
	3 [high]	-	2 [high]	2 [low]	-3.0614	0.329	44.0	-9.307	< .001
		-	2 [high]	3 [high]	-2.9855	0.328	44.0	-9.090	< .001
2 [high]	2 [low]	-	2 [high]	3 [high]	0.0759	0.309	44.0	0.246	1.000

2.3 Discussion

In Experiment 1, increasing the subtask A load and subtask B load decreased performance and increased reaction times only on their respective tasks. Increasing subtask B load from 2 to 3 led to increased subtask B error, recall RT, and presentation RT. Similarly, increasing subtask A load from 1 to 2 led to increased subtask A error, recall RT, and presentation RT. Crucially, increasing the number of subtask A lines did not affect performance on the concurrently executed subtask B, and increasing the number of subtask B lines did not affect performance on the concurrently executed subtask A. To put it in another way, load-effect [i.e. set-size effect] of subtask A was only present on subtask A performance, and load-effect of subtask B was only present on subtask B

performance. Let's summarize these findings by two effects: the presence of within-subtask load-effect [i.e. the load of a subtask affected the performance in the respective subtask] and the lack of across-subtasks load-effect [i.e. the load of a subtasks did not affect the performance in the concurrent subtask]. Any possible explanation should account for these two core findings of Experiment 1.

The presence of within-subtask load effect in Experiment 1 is in line with what is usually suggested by the literature: Memory accuracy decreases with the increased amount of information to be held in memory [38; 37]. This general finding is, in fact, one of the main phenomena that would directly express the presence of WMC in previous literature. In our case, within-subtask load-effect may indicate a WMC for a given subtask's WM information. This capacity-limit for the WM information belonging to a subtask may reflect a resource competition amongst the information that belongs to the same subtask to be maintained or a recall interference amongst the information of the same subtask. Regardless, it implies that the recall precision of a given subtask would decrease with the increased amount of information that belongs to that subtask.

Another direct indication of WM's being capacity limited is cited as the disruptive effect of the distractor tasks on the memory tasks in classical dual-task studies. The equivalent of such an effect in Experiment 1 would be across-subtasks load-effect, and crucially, it is not present. The lack of across-subtasks load-effect in Experiment 1 is at odds with the views suggesting that WM's storage and processing [maintenance and manipulation] functions share a common resource [31]. Similarly, Baddeley's model would also predict

that the interference should have happened among the items belonging to the same subtask and between subtask elements, considering both subtasks would rely on the same ‘slave system,’ the visuospatial sketch pad [16]. However, in Experiment 1, the interference of loads only occurred within subtasks, and not between subtasks. Increasing the number of lines in a subtask decreased the performance in the corresponding subtask, but no such interference is observed across subtasks. The system-based theories or the accounts that see the storage and processing as a single resource cannot account for the findings of Experiment 1.

One alternative to give an account of the lack of across-subtasks load-effect despite the presence of within-subtask load-effect is that perhaps information withheld during the processing of a concurrent task [i.e. Subtask A information during the execution of Subtask B] was maintained in a passive store, while the concurrently processed information [i.e. Subtask B information] was in an active store. This maintenance of withheld WM information may qualify to be ‘passive’ for not requiring cognitive resources or perhaps for being dependent on residual resources that is not currently engaged by the concurrent task [i.e. Subtask B]. Regardless, increasing Subtask A load would not affect the performance on Subtask B since they are practically in different stores, while it would affect the performance on Subtask A due to the reasons addressed before [i.e. resource competition or recall interference within the same store]. Similarly, increasing subtask B load would not affect the performance on subtask A which is held in a passive store during the subtask B execution, while decreasing the performance on the respective subtask B.

The activated part of LTM that holds the information without disrupting the WM operations in the attentional focus in state-based theories of WM may be a derivation of such ‘passive store’. The depiction of aLTM is such that it resides on the residual cognitive resources to keep the information in the background, and in that it is capacity-unlimited [27; 12]. Cowan’s and Oberauer’s WM models can be turned to explain the lack of across-subtasks load effect [27; 12] to varying degrees. Cowan’s embedded process model allows for information in aLTM to be held in the background without distraction from FoA [27; 31; 12]. If seemingly independent storage and processing of WM information in Experiment 1 were explained in these terms, it would be like the following: When subtask A information is being kept during the processing of subtask B information, subtask A information is, in fact, stored in aLTM and the current subtask B information was in FoA. And when subtask B execution is over, it is removed from the FoA, and subtask A information is now brought into the attentional focus. Since aLTM is a capacity-unlimited store, the subtask A load kept in there during the execution of subtask B did not affect the subtask B. Similar explanations for the current findings can also be made from Oberauer’s concentric point of view. Here, however, the subtask A information is kept in the aLTM while the subtask B information was in the direct access region after encoding (mB). During the recall of subtask B (rB), the probed item is selectively brought into the attentional focus or FoA. After the execution of subtask B is over, subtask B information is adaptively removed from the direct access region, and this time subtask A information is brought into direct access from aLTM, and a probed item then was in attentional focus (during rA). More specifically, when a set of oriented lines is

in the direct access region, they are directly accessible to the selection of FoA because they are temporarily bounded to the locations in a mental space (in relation to their spatial positions on the screen), which eventually served as cues to selectively retrieve the probed one into the FoA [81]. Both of these explanations account for the findings that there is a within-subtask load effect or interference for all output measures created by the load of the relevant subtask: because during any execution (rA or rB), the relevant subtask's WM contents were in a capacity-limited portion of WM and that is why choosing the probed one amongst them created interference due to the increased number of items in the set [81]. Cowan's and Oberauer's frameworks can also account for why there is no across-subtask interference of loads on any of the output measures: subtask A information is kept in the background in the aLTM while the subtask B is maintained and executed in the capacity limited part of WM (direct access region in Oberauer, and FoA in Cowan) by the operations of central attention.

One alternative to the 'passive storage' explanation we visited may be that, at least occasionally, WM information of separate concurrent tasks is maintained in their corresponding separate goal-directed task-related programs so that their WM loads do not interfere. When subtask A's information is to be held in mind during the execution of subtask B, the information was in a separate VM allocated only to subtask A and subsumed by subtask A's task-related program so that there was no interference from it to subtask B's performance. Similarly, subtask B information is maintained and executed in subtask B's WM, which is subsumed by subtask B's program.

Both alternatives, namely the passive storage and maintenance of WM information as part of the separate task-related programs, seems to account for the core findings of the presence of a within-subtask load-effect and the lack of across-subtasks load-effect in Experiment 1, and the ambiguity cannot be resolved solely based on the input from Experiment 1. It is essential to note that these two alternatives are different from each other in how the withheld information should be kept separate from the concurrent information for it to not affect the concurrent WM performance. While the maintenance by separate task-related programs indicates a deliberate and goal-directed maintenance of withheld information, the passive store idea would predict a maintenance that cannot be goal-directed, and necessarily subject to the effects of passage of time (decay) and further deployment of cognitive resources. To resolve this ambiguity, we conducted Experiment 2.

Chapter 3

Experiment 2

Experiment 1 showed that increasing the number of lines related to the withheld subtask (or subtask A) only affected the performance on subtask A and not subtask B, suggesting that information related to different subtasks did not interact. Perhaps this was because WM information related to different subtasks can be organized as parts of different programs that did not interfere with each other. Alternatively, it may be claimed that the information related to the withheld subtask A was held in a passive store that either did not require cognitive resources or perhaps depended on residual resources not engaged by subtask B. In either case, increasing subtask A related information would degrade the fidelity of subtask A related recall [due to resource competition or recall interference] but not affect subtask B related recall since A and B lines were in separate stores. During subtask B only B lines were in the active working memory store. A lines at this time were in a passive store e.g. in the residually activated but decaying long term memory. With Experiment 2 we disambiguate between these two possibilities.

Here participants executed long trial episodes that consisted of 4 sequential phases (Figure 3.1). Trial episodes began with the presentation of an 'A' line (lets call it A1 line). As in Experiment 1, participants were to keep its orientation in

mind. This was followed by the presentation of two B lines [mB1]. Participants were then probed to recall the orientation of one of them (rB1). Unlike experiment 1, this was not followed by the recall of A1 line. Instead, participants were presented with a second A line (A2 line). They were to also keep its orientation in mind. Participants then saw a second set of B lines (mB2), followed by the probed recall of the orientation of one of them (rB2). Now, the participants had to recall, not the A2 but, the A1 line orientation (rA1). They were then presented with a third A line (A3). They then saw a third set of B lines (mB3), followed by the probed recall of the orientation of one of them (rB3). This was followed by the recall of the A2 line (rA2). In the last phase, no A lines were presented. Participants saw a fourth set of B lines (mB4), followed by the probed recall of the orientation of one of them (rB4). Lastly, they recalled the orientation of the A3 line (rA3).

Note that this design requires recollection of not the last A line but the one before that. This forces participants to actively organize A lines in their working memory, else the latest A line would interfere with the recollection of the A line presented prior to it. Further, this design requires active removal of the already recalled A lines, else performance would decrease across the four sequential phases. Both of these properties - task related organization of information and removal of information no longer relevant are well-identified features of working memory [13; 12].

Further, the number of A lines being maintained during the recall of A or B lines was variable across the four phases. During B1 and B4, participants had only

one A line to maintain while during B2 and B3, they had two A lines. During the recall of A1 and A2 lines in phases 2 and 3, they were keeping two A line orientations in mind but during the recall of A3 line in phase 4, they had to keep only one A line orientation in mind.

If the number of A lines affected the fidelity of recall of B lines then performance should be poorer on B2 and B3 compared to B1 and B4. However, if A and B lines did not interfere with each other then performance should be the same across B1 to B4.

If A lines were held in a passive store then performance should decline across the recall of A1 and A2 since interference between A lines will start to pile up. They have lines A1 and A2 when they are recalling A1 but by the time they are recalling A2, they have had lines A1, A2 and A3.

The current experiment allows us to test a very specific prediction of the view that information related to different tasks [or subtasks] may not interfere with each other. Recall that in Experiment 1 increasing the number of A lines in memory from one to two decreased the fidelity of their recall. The two A lines in experiment 1 were part of the same subtask. In the current experiment participants were maintaining two A lines when recalling A1 and A2 and only one A line when recalling A3. However, the two A lines being maintained during the recall of A1 (A1 and A2) or A2 (A2 and A3) were part of different subtasks. Thus when recalling A2, participants have lines A2 and A3 in their mind but they knew that A2 and A3 will be recalled at different task junctures. It is plausible that therefore

lines A2 and A3 will be organized as parts of different programs. This would predict that recollection of A1 and A2 lines will not show decreased fidelity compared to the recall of A3 lines even though during the former two A lines are being maintained while only one line is being maintained during the latter. In contrast, every other working memory account would predict that fidelity of recall should be poorer for A1 and A2 compared to A3.

All in all, the current design of Experiment 2 allowed us to see if the finding of lack of across-subtasks load-effect in Experiment 1 is replicable in Experiment 2, and more crucially to test the specific predictions of two alternatives for the findings of Experiment 1, namely the ‘passive storage idea’ and the separate maintenance of information that belongs to different task-related programs (programs-related view):

- a. If A and B lines did not interfere with each other then performance should be the same across B1 to B4.
- b. If the number of A lines affected the recall of B lines then performance should be poorer on rB2 and rB3 (i.e. two A lines are held) compared to rB1 and rB4 (i.e. one A line is held).
- c. If A lines were held in a passive store then performance should decline across the recall of A1 and A2 (i.e. across episode 2 and episode 3) due to interference between A lines that will start to pile up.
- d. If lines A1, A2 and A3 were organized as parts of different programs, this would predict that recollection of A1 and A2 lines will not show decreased fidelity compared to the recall of A3 lines even though during the

former two A lines are being maintained while only one line is being maintained during the latter.

e. Every other working memory account other than A lines' being organized as parts of different programs (including the 'passive store' views) would predict that fidelity of recall should be poorer for A1 and A2 compared to A3.

The aim of Experiment 2 was to see the replicability of the findings of Experiment 1 and to disambiguate between two alternatives for the core findings in Experiment 1 by simply testing their specific predictions mentioned above.

3.1 Methods

3.1.1 Procedure

Experimental procedure was similar to the Experiment 1, except that the reimbursement was done as either course credit or a payment of £25 regardless of the performance.

3.1.2 Participants

54 healthy participants [39 females] within an age group of 19-56 ($M=24.18$, $SD=11.35$) were recruited. Participants are given digital copies of the informed consent and they sign them with their e-signature before participation in the experiment. Participants were reimbursed with course credits or with payment of £25.

3.1.3 Experimental Paradigm

In Experiment 2, similar to Experiment 1, participants memorized the orientations of the lines presented, and recalled them when probed for subtask A and subtask

B, on separate black and pink screens, respectively. However, crucially, in Experiment 2, four episodes of Experiment 1 are organized into a complex 4-part task with the added condition that A lines of a part be recalled not in that part but in the next part. In this design, in the first episode of a session, participants first memorized an A line's orientation (mA1), then memorized two B lines' orientations (mB1), then probed to recall one of them (rB1). The first episode did not have a subtask A recall, since each A line presented was memorized to be recalled not in the immediate episode but in the next episode. In the second episode of a session, participants memorized another A line's orientation (mA2), then memorized two new B lines' orientations (mB2), then probed to recall one of them (rB2), and finally probed to recall the A line that they saw in the first episode (rA1). In the third episode of a session, participants memorized another A line's orientation (mA3), then memorized two new B lines' orientations (mB3), then probed to recall one of them (rB3), and finally probed to recall the A line that they saw in the second episode (rA2). In the fourth and final episode of the session, there was no new A line to be presented, participants saw two new B lines (mB4), probed to recall one of their orientation (rB4), and finally probed to recall the orientation of the A line they saw in the third episode (rA3). The task structure was:

mA1-mB1-rB1—mA2-mB2-rB2-rA1—mA3-mB3-rB3-rA2—mB4-rB4-rA3.

Other details were very similar to Experiment 1 (Figure 3.1).

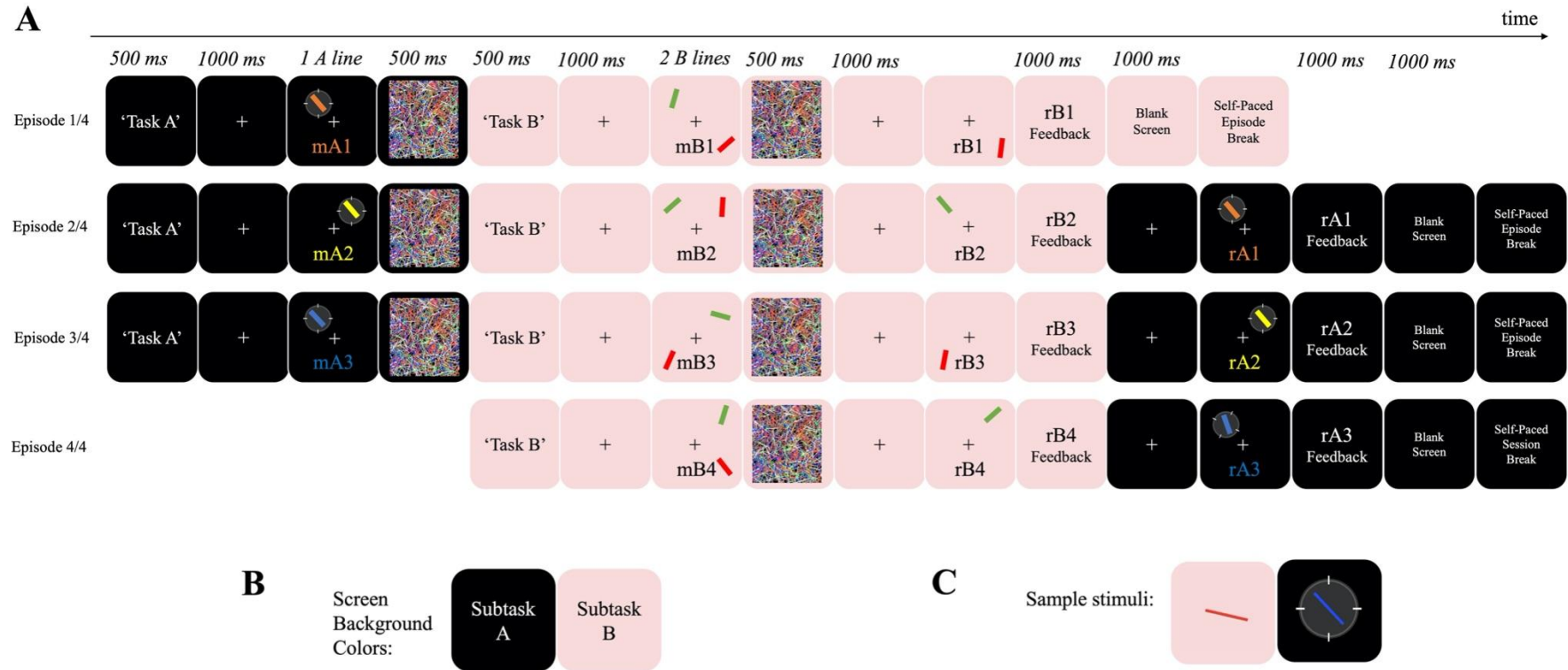
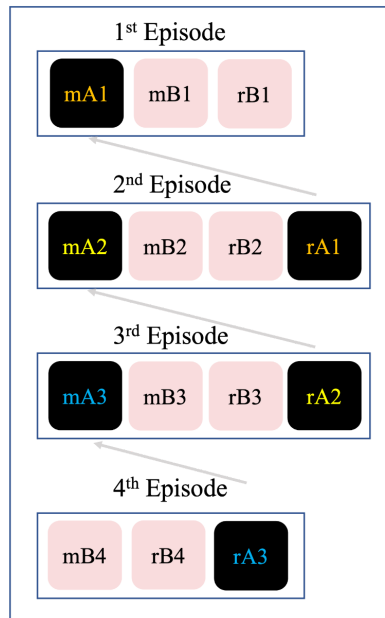


Figure 3.1. The Experimental Paradigm of Experiment 2. (A) The order of events and screens in Experiment 2. A session consisted of four episodes. (B) Different subtasks were indicated by different screen background colors. Subtask A was indicated by black, and subtask B was indicated by pink screen. (C) Relevant stimuli were colored oriented lines. One sample line per background is shown in the figure, they were taken from the memorization phase of the corresponding subtask. Note that the subtask A line was presented with a circle around it, while the subtask B line was not.

Figure 3.2 illustrates the session structure of Experiment 2 (Figure 3.2.a) and the episode structure of other experiments in this thesis (Figure 3.2.b) in a comparative way. Another departure from Experiment 1 was that during each mA one A line was presented, and during each mB two B lines were presented. In a given session, the colors of the lines presented in all mAs had different colors from each other and from B lines.

(a) A Session in Experiment 2



(b) An Episode in Experiment 1, 3, and 4

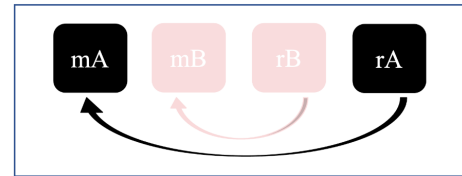


Figure 3.2. Structure of Experiment 2, in comparison with the structure of Experiment 1, 3, and 4. (a) In Experiment 2, the smallest task unit in which the task goals are completed was a session which consisted of four episodes. (b) In Experiments 1, 3 and 4, the smallest task unit in which the task goals are completed was an episode, and each session consisted of four episodes with identical structure.

The experiment consisted of 20 sessions and each session was a 4-episode complex task. The pool of colors for subtask A and subtask B was different. Similarly, the colors of A1, A2, and A3 were chosen from separate pools, as well. Between episodes and sessions, participants could take breaks if they wanted to.

3.1.4. Data Analysis

All analyses were conducted on MATLAB R2021b, and figures illustrating analysis results were created in Prism GraphPad Software [version 9.4.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com] by using the mean and 95% Confidence Interval (CI) information gathered from MATLAB analyses. 95% CIs were calculated by using Cousineau-Morey method for within-participant or repeated measures designs [77; 78]. Descriptive statistics tables and paired-sample t-test tables are gathered from Jamovi [79; 80].

In Experiment 2, as mentioned, the number of lines presented for subtask B was not manipulated, it was always 2. This left us with two different withheld subtask A WM load conditions to compare for any output measure: 1 and 2 lines (or low and high). During a given recall of a subtask B (an rB) participants were holding either 1 or 2 subtask A-lines in their minds. In the rB of episode 1 and episode 4 (rB1 and rB4) participants were holding one A-line, and in the rB of episode 2 and episode 3 (rB2 and rB3) participants were holding two A-lines. Similarly, during any recall of subtask A (an rA) participants were either holding one or two A-lines in their minds, meaning that they were either recalling the only A-line they were supposed to hold in their mind (episode 4 - rA3), or there was another A line in their mind (that belonged to another program) which they will be prompted to recall in the next episode (episode 2 - rA1 and episode 3 - rA2).

Therefore, the number of A lines to be held in mind could be grouped across some episodes if they had the same number of A lines, or performance could have been investigated across all the episodes of a session, for respective output measures. This meant that analyzing the Experiment 2 data was possible in two ways:

1. Analyzing by considering withheld subtask A WM load [i.e. number of A linesheld] to see the effect of the subtask A WM load on subtask A and subtask B performance: *Subtask A WM-load-wise analysis*. This approach is relevant for testing the predictions b, d, and e in the introduction of this thesis. It would require us to aggregate data from relevant episodes and phases to form two groups of withheld subtask A WM loads for a given output measure: subtask A WM load of 1 (i.e. one A line held) and subtask A WM load of 2 (i.e. two A lines held). For example, for subtask B error, we would aggregate data from rB1 and rB4 to construct mean error for subtask A WM load of 1, and from rB2 and rB3 to construct mean error for subtask A WM load of 2. Two groups of subtask A WM loads were to be compared in paired-samples T-test analyses for each output measure.

2. Analyzing the data throughout the episodes of a session to see whether the performance changes across the episodes of a session: *episode-wise analysis*. This approach is relevant for the predictions a, c, and d in the introduction of this chapter. Note that not all episodes consisted of data of all output measures. One-way repeated measures ANOVAs were to be conducted to see whether different episodes are differing from each other on a given output measure. Also,

planned pairwise comparisons were to be performed to compare different episodes.

3.2 Results

3.2.1 Subtask A WM Load-wise Analyses

Figure 3.3 illustrates the paired samples T-test results for subtask A WM load-wise comparisons for Experiment 2.

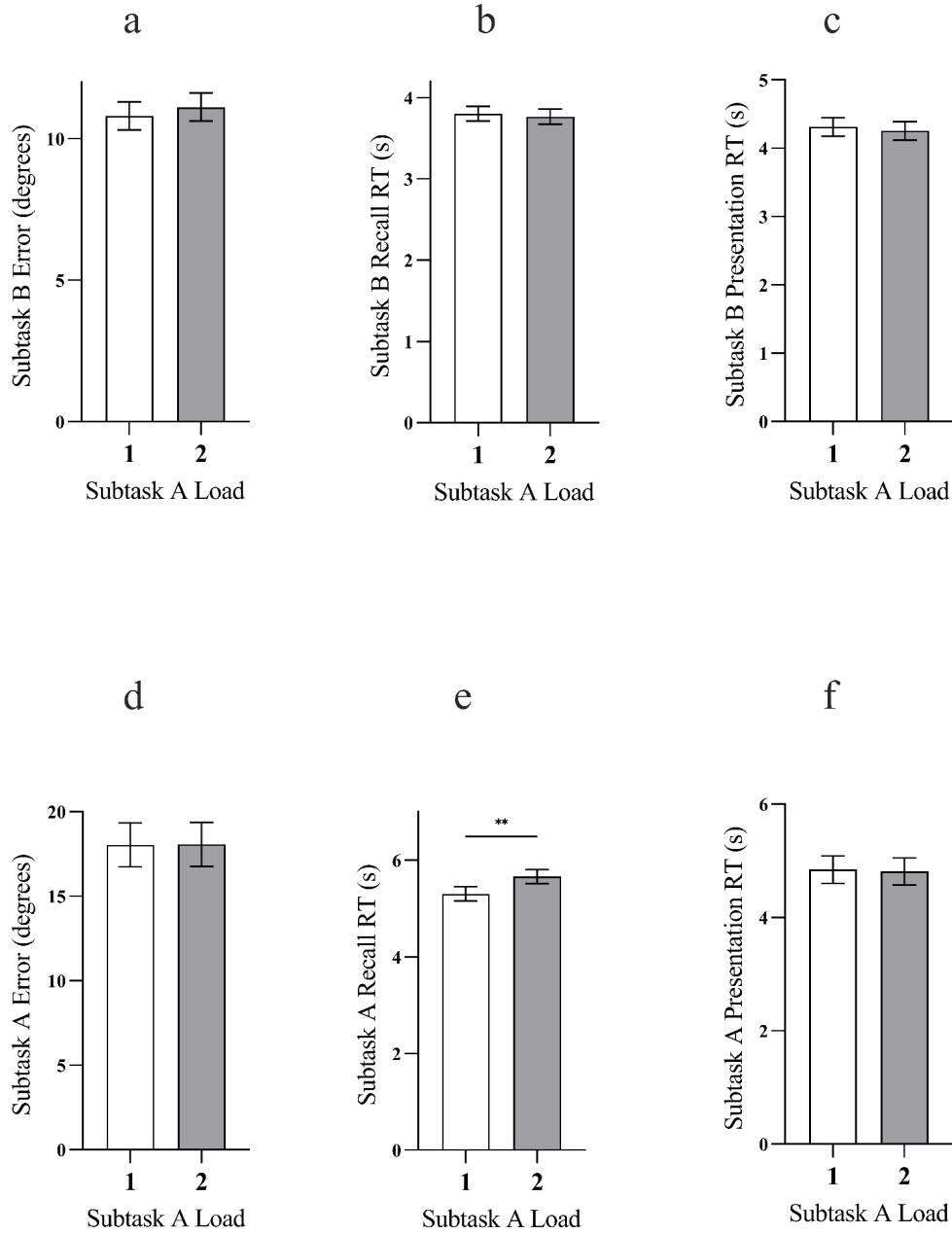


Figure 3.3. Paired Samples T-test Results for Subtask A WM Load-wise Comparisons for Experiment 2. Output measures are: (a) subtask B error [degrees], (b) subtask B recall RT (s), (c) subtask B presentation RT (s), (d) subtask A error (degrees), (e) subtask A recall RT (s), (f) subtask A presentation RT (s). Error bars are 95% CIs of the individual means.

3.2.1.1 Subtask B Error

Table 3.1. Descriptives for Subtask B Error [degrees] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	54	54
Missing	0	0
Mean	10.8	11.1
Std. error mean	0.872	0.952
95% CI mean lower bound	9.10	9.25
95% CI mean upper bound	12.5	13.0
Median	9.21	8.95
Mode	7.03 ^a	6.50 ^a
Standard deviation	6.41	7.00
Minimum	2.83	3.00
Maximum	32.3	39.0
Skewness	1.64	2.22
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.2. Paired Samples T-test for Subtask B Error [degrees] in Subtask A Load-wise Analysis in Experiment 2

		statistic	df	p	Mean difference	SE difference	95% Confidence Interval			Effect Size	95% Confidence Interval				
							Lower	Upper	Lower		Upper				
AL: 1 [low]	AL: 2 [high]						Student's t								
		-0.89	53	0.37	-0.31	0.348	-1.01	0.38	Cohen's d	-0.12	-0.39	0.146			

Subtask B error was calculated as the mean deviation of the orientation recalled in rB from the target B line that was presented in the same episode's mB phase. Table 3.1. shows the descriptives of subtask A load conditions for subtask B error. Mean deviations of Episode 1 (rB1) and Episode 4 (rB4) were calculated to find the subtask B error for the condition where withheld subtask A WM load was 1 ($M = 10.8$, $SD = 6.41$). Mean deviations of Episode 2 (rB2) and Episode 3 (rB3) were calculated to find the subtask B error for the condition where 2 A lines were held in mind ($M = 11.1$, $SD = 7.00$). The paired samples t-test was conducted to compare the subtask B error when 1 A line vs 2 A lines are held. It is revealed that the group means did not differ significantly ($t(53) = 0.90$, $p = 0.37$, $CI =$

[-1.01, 0.39], $d = -0.122$), the number of A lines held in mind did not affect subtask B error. Results are illustrated in Figure 3.3.a.

3.2.1.2 Subtask B Recall RT

Table 3.3. Descriptives for Subtask B Recall RT [s] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1	Subtask A Load: 2
N	54	54
Missing	0	0
Mean	3.80	3.77
Std. error mean	0.240	0.243
95% CI mean lower bound	3.33	3.29
95% CI mean upper bound	4.27	4.24
Median	3.43	3.29
Mode	1.75 ^a	1.34 ^a
Standard deviation	1.77	1.79
Minimum	1.75	1.34
Maximum	10.7	10.0
Skewness	1.83	1.43
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.4. Paired Samples T-test for Subtask B Recall RT [s] in Subtask A Load-wise Analysis in Experiment 2

										95% Confidence Interval		95% Confidence Interval	
			statistic	df	p	Mean difference	SE difference	Lower	Upper		Effect Size	Lower	Upper
AL: 1	AL: 2	Student's t	0.587	53.0	0.560	0.0380	0.0648	-0.0919	0.168	Cohen's d	0.0798	-0.188	0.347

Subtask B recall RT is calculated as the mean time participants took to answer the probe in the rB phase of an episode. Table 3.3. shows the descriptives of subtask A load conditions for subtask B recall RT. Mean reaction time of Episode 1 and Episode 4 recalls (rB1 and rB4) was calculated to find the mean subtask B recall RT for the condition where subtask A WM load was 1 ($M = 3.80$, $SD = 1.77$). Mean reaction time of Episode 2 and Episode 3 recalls (rB2 and rB3) was calculated to find the subtask B recall RT for the condition where withheld

subtask A WM load was 2 ($M = 2.77$, $SD = 1.79$). A paired samples t-test was conducted to compare the subtask B recall RT when subtask A load was 1 and when subtask A load was 2. It was revealed that the group means did not differ significantly ($t(53) = 0.59$, $p = 0.56$, $CI\ 95\% = [-0.09, 0.17]$, $d = 0.08$). The subtask A load or the number of A lines held in mind did not affect concurrently processed subtask B's recall RT. Results are illustrated in Figure 3.3.b.

3.2.1.3 Subtask B Presentation RT

Table 3.5. Descriptives for Subtask B Presentation RT [s] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	54	54
Missing	0	0
Mean	4.31	4.25
Std. error mean	0.313	0.338
95% CI mean lower bound	3.70	3.59
95% CI mean upper bound	4.93	4.92
Median	4.18	3.77
Mode	0.939 ^a	0.910 ^a
Standard deviation	2.30	2.48
Minimum	0.939	0.910
Maximum	12.0	12.1
Skewness	0.908	0.945
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.6. Paired Samples T-test for Subtask B Presentation RT [s] in Subtask A Load-wise Analysis in Experiment 2

Experiment 2													
		statistic	df	p	Mean difference	SE difference	95% Confidence Interval			Effect Size	95% Confidence Interval		
							Lower	Upper	Lower		Upper		
AL: 1 [low]	AL: 2 [high]	Student's t	0.623	53.0	0.536	0.0589	0.0946	-0.131	0.249	Cohen's d	0.0847	-0.183	0.352

Subtask B presentation RT is calculated as the mean time participants took to memorize the B lines in mB phase of a respective episode. Table 3.5. shows the descriptives of two subtask A load conditions for subtask B presentation RT which were, again, A load of 1 (low) and A load of 2 (high). Mean presentation time of Episode 1 and Episode 4 mBs (mB1 and mB4) was calculated for the condition where subtask A WM load was 1 ($M = 4.31$, $SD = 2.30$). Mean presentation time of Episode 2 and Episode 3 mBs [mB2 and mB3] was calculated for the condition where withheld subtask A WM load was 2 ($M = 4.25$, $SD = 2.48$). A paired samples T-test was conducted to compare the subtask B presentation RTs of subtask A load was low and high. It was revealed that the group means did not differ significantly ($t(53) = 0.62$, $p = 0.53$, $CI\ 95\% = [-0.13, 0.25]$, $d = 0.08$). The subtask A load did not affect the subtask B presentation RT. Results are illustrated in Figure 3.3.c.

3.2.1.4 Subtask A Error

Table 3.7. Descriptives for Subtask A Error [degrees] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1	Subtask A Load: 2
N	54	54
Missing	0	0
Mean	18.0	18.1
Std. error mean	1.87	1.84
95% CI mean lower bound	14.4	14.5
95% CI mean upper bound	21.7	21.7
Median	13.7	14.1
Mode	2.19 ^a	14.1 ^a
Standard deviation	13.7	13.5
Minimum	2.19	2.17
Maximum	53.1	56.8
Skewness	0.833	0.924
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.8. Paired Samples T-test for Subtask A Error [degrees] in Subtask A Load-wise Analysis in Experiment 2

		statistic	df	p	Mean difference	SE difference	95% Confidence Interval		Cohen's d	Effect Size	95% Confidence Interval	
AL: 1	AL: 2						Lower	Upper			Lower	Upper
							Student's t					
[low]	[high]	-0.0264	53.0	0.979	-0.0242	0.918	-1.86	1.82		-0.00359	-0.270	0.263

Subtask A error was calculated as the mean deviation of the orientation recalled in rA from the target A line that was presented in the previous episode's mA phase. Table 3.7. shows the descriptives of subtask A load conditions for subtask A error. Mean deviation of recall of Episode 2 (rA1) and Episode 3 (rA2) were calculated for the condition where withheld subtask A WM load was 2 ($M = 18.1$, $SD = 13.5$), and compared to the mean deviation of Episode 4 (rA3), which had a subtask A WM load of 1 ($M = 18.0$, $SD = 13.7$). Paired samples T-test was conducted to compare the subtask A error in these groups. It was revealed that the groups did not significantly differ from each other ($t(53) = -0.02$, $p = 0.98$, $CI\ 95\% = [-1.86, 1.82]$, $d = -0.004$). Although rA1 and rA2 had 2 A lines to hold, it did not result in a poorer accuracy for subtask A recall compared to rA3 where only one A line was held in mind. Results are illustrated in Figure 3.3.d.

3.2.1.4 Subtask A Recall RT

Table 3.9. Descriptives for Subtask A Recall RT [s] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1	Subtask A Load: 2
N	54	54
Missing	0	0
Mean	5.30	5.66
Std. error mean	0.294	0.314
95% CI mean lower bound	4.72	5.05
95% CI mean upper bound	5.88	6.28
Median	4.91	5.11
Mode	2.29 ^a	2.23 ^a
Standard deviation	2.16	2.30
Minimum	2.29	2.23
Maximum	12.2	12.2

Table 3.9. Descriptives for Subtask A Recall RT [s] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1	Subtask A Load: 2
Skewness	0.904	0.896
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.10. Paired Samples T-test for Subtask A Recall RT [s] in Subtask A Load-wise Analysis in Experiment 2

										95% Confidence Interval		95% Confidence Interval	
			statistic	df	p	Mean difference	SE difference	Lower	Upper		Effect Size	Lower	Upper
AL: 1 [low]	AL: 2 [high]	Student's t	-3.46	53.0	0.001	-0.360	0.104	-0.569	-0.152	Cohen's d	-0.471	-0.751	-0.188

Subtask A recall RT is calculated as the mean time participants took to answer to the probe in the rA phase of a respective episode. Table 3.9. shows the descriptives of subtask A load conditions for subtask A recall RT. Mean recall RT of Episode 2 (rA1) and Episode 3 (rA2) were calculated for the condition where withheld subtasks A WM load was 2 ($M = 5.66$, $SD = 2.30$), and compared to the mean recall RT of Episode 4 (rA3), which had a withheld subtask A WM load of 1 ($M = 5.30$, $SD = 2.16$). Paired samples t-test was conducted to compare the subtask A recall RTs in these groups. It was revealed that the groups significantly differed from each other ($t(53) = -3.46$, $p = 0.001$, $CI\ 95\% = [-0.57, 0.15]$, $d = -0.47$). The additional subtask A load that is held in relation to another task related program increased the reaction time in a concurrent subtask A. Results are illustrated in Figure 3.3.e.

3.2.1.6 Subtask A Presentation RT

Table 3.11. Descriptives for Subtask A Presentation RT [s] in Subtask A Load-wise Analysis in Experiment 2

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	54	54
Missing	0	0
Mean	4.85	4.81
Std. error mean	0.324	0.334
95% CI mean lower bound	4.21	4.16
95% CI mean upper bound	5.48	5.47
Median	4.22	4.55
Mode	1.43 ^a	1.59 ^a
Standard deviation	2.38	2.46
Minimum	1.43	1.59
Maximum	11.4	13.4
Skewness	0.870	1.46
Std. error skewness	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.12. Paired Samples T-test for Subtask A Presentation RT [s] in Subtask A Load-wise Analysis in Experiment 2

										95% Confidence Interval		95% Confidence Interval		
										Lower	Upper	Effect Size	Lower	Upper
AL: 1 [low]	AL: 2 [high]	Student's t	0.184	53.0	0.855	0.0311	0.169	-0.309	0.371	Cohen's d	0.0250	-0.242	0.292	

Subtask A presentation RT is calculated as the mean time participants took to memorize the A line in mA phase of a respective episode. Table 3.11. shows the descriptives of subtask A load conditions for subtask A presentation RT. Mean presentation time of Episode 1 mA (mA1) was calculated for the condition where subtask A WM load was 1 ($M = 4.85$, $SD = 2.38$). Mean presentation time of Episode 2 (mA2) and Episode 3 (mA3) were calculated for the condition where withheld subtask A WM load was 2 ($M = 4.81$, $SD = 2.46$). The paired samples t-test was conducted to compare the subtask A presentation RT when subtask A load was 1 and when subtask A load was 2. It was revealed that the group means did not differ significantly ($t(53) = 0.184$, $p = 0.86$, $CI\ 95\% = [-0.31, 0.37]$, $d = 0.03$). Holding an additional A line in mind did not increase the RT in mA2 and mA3 compared to mA1. Results are illustrated in Figure 3.3.f.

3.2.2 Episode-wise Analyses

Episode-wise analyses were complementary to the subtask A WM load-wise analyses in the way that it would allow us to see the performance across the episodes of a larger session. Figure 3.4 illustrates the results of one-way repeated measures ANOVA and planned comparison results for episode-wise analyses in Experiment 2.

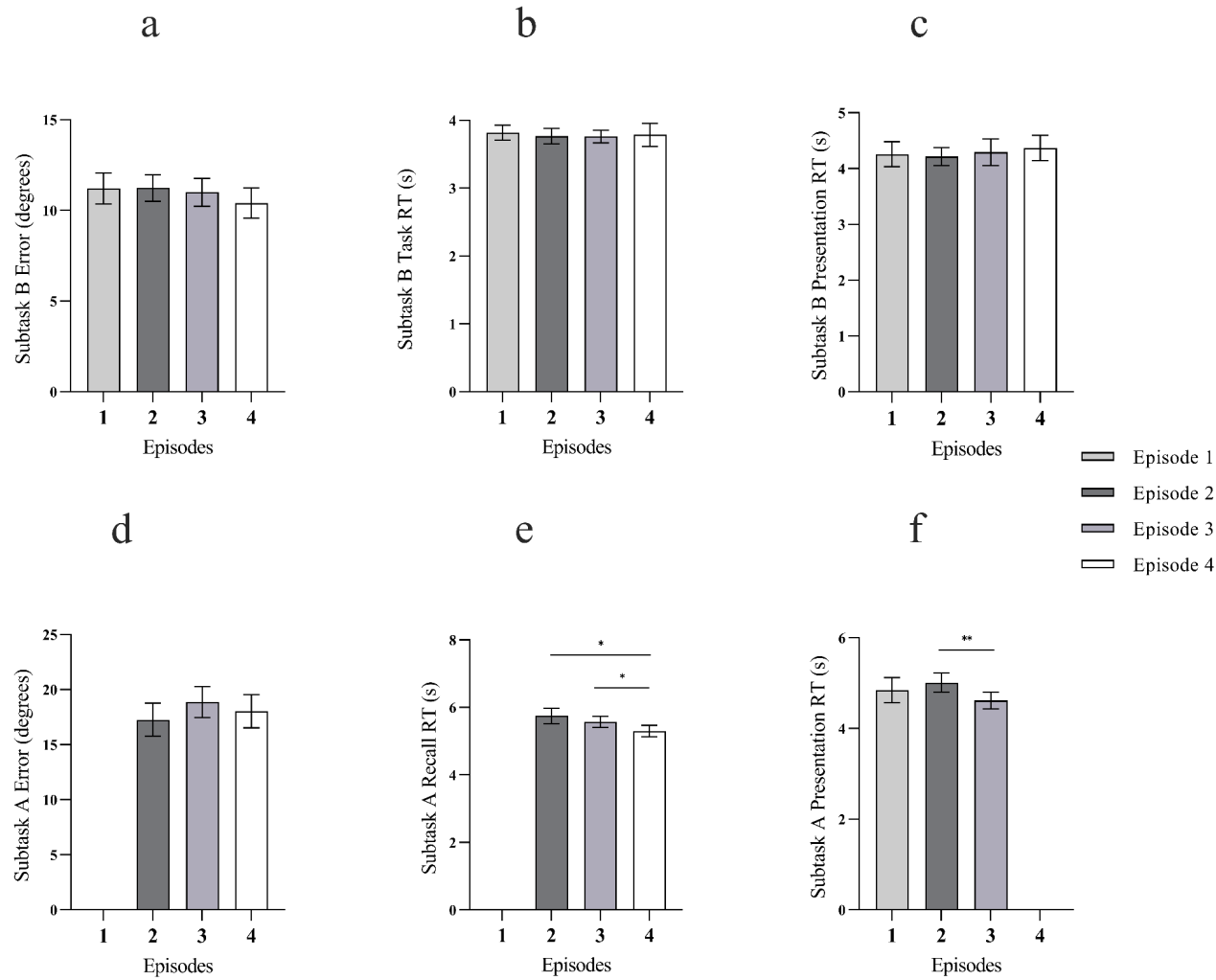


Figure 3.4. One-way Repeated Measures ANOVA and Planned Comparison Results for Episode-wise Analyses in Experiment 2. Output measures are: (a) subtask B error (degrees), (b) subtask B recall RT (s), (c) subtask B presentation RT (s), (d) subtask A error (degrees), (e) subtask A recall RT (s), (f) subtask A presentation RT (s). Error bars represent 95% CIs.

3.2.2.1 Subtask B Error

Subtask B error for a given episode is calculated as the mean deviation of the recall in rB from the target shown in respective mB of the same episode, in degrees. All episodes in a given session of Experiment 2 had mB and rB phases, and that left us with four episodes' subtask B errors to be compared (see the descriptives in Table 3.13) in a one-way repeated measures ANOVA with four levels (episode 1, episode 2, episode 3, and episode 4). Within-subjects effects are shown in Table 3.14. There was not a statistically significant difference amongst these four groups ($F(3,159) = 0.94$, $p = 0.42$, $\eta_p^2 = 0.017$). Planned pairwise comparisons with Bonferroni correction (Table 3.15) revealed no significant differences amongst any pairs of episode mean scores for subtask B error. There were two A lines to be held in mind during the subtask B is execution of episode 2 (rB2) and episode 3 (rB3). Recalls of subtask B in episode 1 (rB1) and episode 4 (rB4), on the other hand, were executed while there was only 1 A line held. Results showed that subtask B accuracy was not poorer for the subtask B execution where there were 2 A lines held concurrently compared to 1 A line, and the subtask B recall RT was the same across the episodes of a session, performance was the same across B1 to B4. The results are illustrated in Figure 3.4.a.

Table 3.13. Descriptives for Subtask B Error [degrees] in Episode-wise Analyses in Experiment 2

	episode1	episode2	episode3	episode4
N	54	54	54	54
Missing	0	0	0	0
Mean	11.2	11.2	11.0	10.4
Std. error mean	0.985	0.969	1.01	0.867
95% CI mean lower bound	9.27	9.34	9.02	8.71
95% CI mean upper bound	13.1	13.1	13.0	12.1
Median	8.30	9.95	8.52	8.63
Mode	6.60	14.2	7.05 ^a	6.43 ^a
Standard deviation	7.24	7.12	7.43	6.37
Minimum	2.40	3.10	2.90	3.10

Table 3.13. Descriptives for Subtask B Error [degrees] in Episode-wise Analyses in Experiment 2

	episode1	episode2	episode3	episode4
Maximum	35.4	41.7	36.4	35.9
Skewness	1.52	2.41	1.92	2.08
Std. error skewness	0.325	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.14. Within Subjects Effects for Subtask B Error [degrees] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	23.8	3	7.95	0.941	0.422	0.002	0.002	0.017
Residual	1343.5	159	8.45					

Note. Type 3 Sums of Squares

Table 3.15. Planned Comparisons -Episodes - for Subtask B Error [degrees] in Episode-wise Analyses in Experiment 2

Comparison		Mean Difference	SE	df	t	p _{bonferroni}
Episodes	Episodes					
episode1	- episode2	-0.0351	0.565	53.0	-0.0621	1.000
	- episode3	0.2041	0.537	53.0	0.3800	1.000
	- episode4	0.7941	0.633	53.0	1.2542	1.000
episode2	- episode3	0.2392	0.544	53.0	0.4401	1.000
	- episode4	0.8292	0.503	53.0	1.6482	0.631
episode3	- episode4	0.5900	0.566	53.0	1.0417	1.000

3.2.2.2 Subtask B Recall RT

Subtask B recall RT for a given episode is calculated as the time participants took in rB to answer the probe, in seconds. All episodes in a given session of Experiment 2 had mB and rB phases, and that left us with four episodes' subtask B recall RTs to be compared (see the descriptives in Table 3.1) in a one-way repeated measures ANOVA with four levels (episode 1, episode 2, episode 3, and episode 4). Within subjects effects are shown in Table 3.17. There was not a statistically significant difference amongst these four groups ($F(3,159) = 0.17, p = 0.92, \eta^2_p = 0.003$). Planned pairwise comparisons with Bonferroni correction (Table 3.18] revealed no significant differences amongst any pairs of episode

mean scores for subtask B recall RT. The recall performance was the same across B1 to B4 for recall RT, too. The results are illustrated in Figure 3.4.b.

Table 3.16. Descriptives for Subtask B Recall RT [s] in Episode-wise Analyses in Experiment 2

	episode1	episode2	episode3	episode4
N	54	54	54	54
Missing	0	0	0	0
Mean	3.82	3.77	3.76	3.79
Std. error mean	0.245	0.251	0.238	0.248
95% CI mean lower bound	3.34	3.27	3.30	3.30
95% CI mean upper bound	4.30	4.26	4.23	4.27
Median	3.26	3.18	3.45	3.34
Mode	1.46 ^a	1.47 ^a	1.22 ^a	1.68 ^a
Standard deviation	1.80	1.85	1.75	1.82
Minimum	1.46	1.47	1.22	1.68
Maximum	10.0	10.3	9.79	11.4
Skewness	1.62	1.44	1.39	1.91
Std. error skewness	0.325	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.17. Within Subjects Effects for Subtask B Recall RT [s] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	0.106	3	0.0353	0.166	0.919	0.000	0.000	0.003
Residual	33.889	159	0.2131					

Note. Type 3 Sums of Squares

Table 3.18. Planned Comparisons -Episodes - for Subtask B Recall RT [s] in Episode-wise Analyses in Experiment 2

Comparison						
Episodes	Episodes	Mean Difference	SE	df	t	P _{bonferroni}
episode1	- episode2	0.05173	0.0771	53.0	0.6708	1.000
	- episode3	0.05618	0.0620	53.0	0.9056	1.000
	- episode4	0.03190	0.1071	53.0	0.2980	1.000
episode2	- episode3	0.00445	0.0619	53.0	0.0720	1.000
	- episode4	-0.01983	0.1079	53.0	-0.1838	1.000
episode3	- episode4	-0.02428	0.1032	53.0	-0.2353	1.000

3.2.2.3 Subtask B Presentation RT

Subtask B presentation RT for a given episode is calculated as the mean of the time participants took to memorize the B-lines in mB, in degrees. All episodes in a given session of Experiment 2 had mB phases, and that left us with four episodes' subtask B recall RTs to be compared (see the descriptives in Table 3.19) in a one-way repeated measures ANOVA with four levels (episode 1, episode 2, episode 3, and episode 4). Within subjects effects are shown in Table 3.20. There was not a statistically significant difference amongst these four groups ($F(3,159) = 0.94$, $p = 0.42$, $\eta^2_p = 0.017$). Planned pairwise comparisons with Bonferroni correction (Table 3.21) revealed no significant differences amongst any pairs of episode mean scores for subtask B presentation RT. Results indicated that the subtask B performance was the same across B1 to B4 in terms of memorization or presentation RT, just like other output measures of subtask B performance. The results are illustrated in Figure 3.3.c.

Table 3.19. Descriptives for Subtask B Presentation RT [s] in Episode-wise Analyses in Experiment 2

	episode1	episode2	episode3	episode4
N	54	54	54	54
Missing	0	0	0	0
Mean	4.26	4.22	4.29	4.37
Std. error mean	0.323	0.329	0.362	0.326
95% CI mean lower bound	3.62	3.57	3.58	3.73
95% CI mean upper bound	4.89	4.86	5.00	5.01
Median	4.38	3.84	3.85	3.75
Mode	0.934 ^a	0.975 ^a	0.846 ^a	0.944 ^a
Standard deviation	2.37	2.42	2.66	2.39
Minimum	0.934	0.975	0.846	0.944
Maximum	13.8	10.9	13.3	10.1
Skewness	1.37	0.854	1.07	0.736
Std. error skewness	0.325	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.20. Within Subjects Effects for Subtask B Presentation RT [s] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	0.681	3	0.227	0.369	0.775	0.001	0.001	0.007
Residual	97.741	159	0.615					

Note. Type 3 Sums of Squares

Table 3.21. Planned Comparisons -Episodes - for Subtask B Presentation RT [s] in Episode-wise Analyses in Experiment 2

Comparison							
Episodes	Episodes	Mean Difference	SE	df	t	$p_{\text{Bonferroni}}$	
episode1	- episode2	0.0418	0.125	53.0	0.335	1.000	
	- episode3	-0.0352	0.165	53.0	-0.214	1.000	
	- episode4	-0.1111	0.168	53.0	-0.661	1.000	
episode2	- episode3	-0.0770	0.149	53.0	-0.517	1.000	
	- episode4	-0.1529	0.125	53.0	-1.220	1.000	
episode3	- episode4	-0.0760	0.167	53.0	-0.455	1.000	

3.2.2.4 Subtask A Error

The subtask A error for a given episode is calculated as the mean deviation of the recall in rA from the target shown in respective mA, in degrees. Only the last three episodes of a given session of Experiment 2 had rA phases, and that left us with three episodes' subtask A errors to be compared (see the descriptives in Table 3.22) in a one-way repeated measures ANOVA with three levels (episode 2, episode 3, and episode 4). Within subjects effects are shown in Table 3.23. There was not a statistically significant difference amongst these three groups ($F(3,159) = 1.18$, $p = 0.31$, $\eta^2_p = 0.022$). Planned pairwise comparisons with Bonferroni correction (Table 3.24) revealed no significant differences amongst any pairs of episode mean scores for subtask A error. Results indicated that subtask A accuracy did not decline across the recall of A1 and A2 (i.e. across episode and episode 3) even though the A lines held are piled up, instead the subtask A recall

accuracy was the same across rA1 to rA3. The results are illustrated in Figure 3.4.d.

Table 3.22. Descriptives for Subtask A Error [degrees] in Episode-wise Analyses in Experiment 2

	episode2	episode3	episode4
N	54	54	54
Missing	0	0	0
Mean	17.3	18.9	18.0
Std. error mean	1.82	1.98	1.87
95% CI mean lower bound	13.7	15.0	14.4
95% CI mean upper bound	20.8	22.8	21.7
Median	12.4	14.3	13.7
Mode	5.20 ^a	2.86 ^a	2.19 ^a
Standard deviation	13.4	14.6	13.7
Minimum	2.19	1.85	2.19
Maximum	51.8	61.9	53.1
Skewness	0.996	0.830	0.833
Std. error skewness	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.23. Within Subjects Effects for Subtask A Error [degrees] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	68.6	2	34.3	1.18	0.312	0.002	0.002	0.022
Residual	3084.1	106	29.1					

Note. Type 3 Sums of Squares

Table 3.24. Planned Comparisons -Episodes - for Subtask A Error [degrees] in Episode-wise Analyses in Experiment 2

Comparison			Mean Difference	SE	df	t	p _{bonferroni}
Episodes	Episodes						
episode2	-	episode3	-1.593	1.02	53.0	-1.568	0.368
	-	episode4	-0.773	1.09	53.0	-0.708	1.000
episode3	-	episode4	0.821	1.01	53.0	0.817	1.000

3.2.2.5 Subtask A Recall RT

The subtask A recall RT for a given episode is calculated as the mean amount of time that it took for participants to answer the probe in rA. Only the last three

episodes of a given session of Experiment 2 had rA phases, and that left us with three episodes' subtask A recall RTs to be compared (rA1, rA2, and rA3) (see the descriptives in Table 3.25) in a one-way repeated measures ANOVA with three levels (episode 2, episode 3, and episode 4). Within-subjects effects are shown in Table 3.26. There was a statistically significant difference amongst these three groups ($F(3,159) = 5.64, p = 0.005, \eta^2_p = 0.096$). Planned pairwise comparisons with Bonferroni correction (Table 3.27) revealed that episode 2 subtask A recall RT (rA1) ($M = 5.75, SD = 2.41$) is significantly higher than episode 4 subtask A recall RT (rA3) ($M = 5.30, SD = 2.16$) ($t(53) = 2.99, p = 0.013$), and episode 3 subtask A recall RT (rA2) ($M = 5.57, SD = 2.33$) is significantly higher than episode 4 subtask A recall RT (rA3) ($t(53) = 2.73, p = 0.026$). Note that rA1 and rA2 had 2 lines to hold where rA3 had 1 A line. The subtask A recall RT for those was $rA1 > rA3$, $rA2 > rA3$, and $rA1 = rA2$. That means during the subtask A recalls where there was an additional A line in mind, the recall RT was higher. The results are illustrated in Figure 3.4.e.

Table 3.25. Descriptives for Subtask A Recall RT [s] in Episode-wise Analyses in Experiment 2

	episode2	episode3	episode4
N	54	54	54
Missing	0	0	0
Mean	5.75	5.57	5.30
Std. error mean	0.327	0.317	0.294
95% CI mean lower bound	5.11	4.95	4.72
95% CI mean upper bound	6.39	6.19	5.88
Median	5.29	4.84	4.91
Mode	2.10 ^a	2.08 ^a	2.29 ^a
Standard deviation	2.41	2.33	2.16
Minimum	2.10	2.08	2.29
Maximum	12.4	12.8	12.2
Skewness	0.768	0.985	0.904
Std. error skewness	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.26. Within Subjects Effects for Subtask A Recall RT [s] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	5.51	2	2.754	5.64	0.005	0.006	0.006	0.096
Residual	51.79	106	0.489					

Note. Type 3 Sums of Squares

Table 3.27. Planned Comparisons -Episodes - for Subtask A Recall RT [s] in Episode-wise Analyses in Experiment 2

Comparison							
Episodes	Episodes	Mean Difference	SE	df	t	$p_{\text{Bonferroni}}$	
episode2	- episode3	0.176	0.1475	53.0	1.19	0.717	
	- episode4	0.448	0.1501	53.0	2.99	0.013	
episode3	- episode4	0.272	0.1000	53.0	2.73	0.026	

3.2.2.6 Subtask A Presentation RT

The subtask A presentation RT for a given episode is calculated as the mean amount of time that it took for participants to memorize the subtask A line presented in a mA. Only the first three episodes of a given session of Experiment 2 had mA phases, and that left us with three episodes' subtask A recall RTs to be compared (mA1, mA2, mA3) (see the descriptives in Table 3.28) in a one-way repeated measures ANOVA with three levels (episode 1, episode 2, and episode 3). Within-subjects effects are shown in Table 3.29. The analysis revealed a marginally non-significant p-value ($F(3,159) = 3.07$, $p = 0.05$, $\eta^2_p = 0.055$). Planned pairwise comparisons with Bonferroni correction (Table 3.30) revealed that episode 2 subtask A presentation RT ($M = 5.01$, $SD = 2.60$) is significantly higher than episode 3 subtask A presentation RT ($M = 4.62$, $SD = 2.38$) ($t(53) = 2.99$, $p = 0.013$). Results indicated that subtask A presentation RT was the same across A1 to A2. And, for A2 the presentation time was higher than A3. The results are illustrated in Figure 3.4.f.

Table 3.28. Descriptives for Subtask A Presentation RT [s] in Episode-wise Analyses in Experiment 2

	episode1	episode2	episode3
N	54	54	54
Missing	0	0	0
Mean	4.85	5.01	4.62
Std. error mean	0.324	0.353	0.324
95% CI mean lower bound	4.21	4.32	3.98
95% CI mean upper bound	5.48	5.71	5.25
Median	4.22	4.60	4.28
Mode	1.43 ^a	1.46 ^a	1.58 ^a
Standard deviation	2.38	2.60	2.38
Minimum	1.43	1.46	1.58
Maximum	11.4	13.9	12.9
Skewness	0.870	1.34	1.52
Std. error skewness	0.325	0.325	0.325

^a More than one mode exists, only the first is reported

Table 3.29. Within Subjects Effects for Subtask A Presentation RT [s] in Episode-wise Analyses in Experiment 2

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Episodes	4.28	2	2.138	3.07	0.050	0.004	0.004	0.055
Residual	73.76	106	0.696					

Note. Type 3 Sums of Squares

Table 3.30. Planned Comparisons -Episodes - for Subtask A Presentation RT [s] in Episode-wise Analyses in Experiment 2

Comparison		Mean Difference	SE	df	t	p _{bonferroni}
Episodes	Episodes					
episode1	- episode2	-0.167	0.186	53.0	-0.899	1.000
	- episode3	0.229	0.172	53.0	1.333	0.564
episode2	- episode3	0.396	0.115	53.0	3.443	0.003

3.3 Discussion

Experiment 2 aimed at disambiguating between the views of ‘passive store’ and WM information of different subtasks’ being organized as parts of separate task-related programs. To achieve that we tested the predictions from these two alternatives.

Let us first visit the results from the tests that were conducted to see whether A lines and B lines interfere in Experiment 2. If A and B lines did not interfere with each other then across B1 to B4 the performance should have been the same for subtask B even though different episodes required to hold different numbers of A lines [i.e. prediction a]. On the other hand, if A and B lines did interfere with each other, when the episodes were grouped in accordance with their subtask A load and compared, the high subtask A load episodes should have led to poorer subtask B performance compared to the low subtask A load episodes [i.e. prediction b]. Episode-wise analyses performed on subtask B related output measures revealed that the subtask B performance was the same across B1 to B4 in terms of recall accuracy, recall RT, and presentation RT, which was in line with the prediction a, suggesting a lack of interference between A and B lines. In addition, the evidence showed that prediction b did not hold. When subtask A load-wise analyses are conducted on subtask B performance measures, low and high load conditions [i.e. 1 A line held vs. 2 A line held] was the same for all output measures of subtask B performance. When these results are taken together, they indicate that the number of A lines held in mind did not affect subtask B performance, hence A lines and B lines did not interfere with each other. This interference-free execution of subtask A lines and subtask B lines replicates the finding of the lack of across-subtask load-effect from Experiment 1.

Now, I will summarize the results regarding the predictions from the ‘passive store’ view. ‘Passive store’ view would predict that performance of subtask A should decline across the recall of A1 and A2 (i.e. across episode 2 and episode 3)

since every added A line would cause increased interference (i.e. prediction c). To test this prediction we compared the subtask A recall performance across relevant episodes. Contrary to this prediction of the ‘passive store’ view, subtask A recall accuracy, and recall RT did not decline across the recall of A1 and A2 (i.e. across episode 1 and episode 3) even though the A lines held are piled up. Even if we extend this prediction c not only to recall but also to memorization performance of subtask A, prediction still did not hold. In terms of the memorization, across the first three episodes, participants saw a new A line in each episode, meaning that the presented A lines have piled up. From the passive store view it would mean that memorization performance should decline across A1 to A3 (presentation RT should increase across A1 to A3). However, the results indicated that presentation RTs for A1 and A2 were the same, and both were significantly higher than A3. All in all, passive store view’s prediction c regarding the subtask A performance did not get validated in Experiment 2. Let us move to the prediction f, which was stated to be that when the fidelity of recall is compared for low and high load conditions (i.e. A3 vs A1 and A2), every WM theory that we visited in the introduction of this thesis, including the passive store views, and other than the view that suggests A lines might be organized as parts of different programs, would predict that the recall should be poorer for high A load conditions (i.e. A1 and A2) compared to low A load condition (i.e. A3) for subtask A performance. However, load-wise comparisons of subtask A recall accuracy did not show any difference between the high and low load conditions. For recall RT, on the other hand, high load had higher RT than low load. This recall RT related result will be visited again later on in this discussion.

Finally, I shall summarize the findings regarding the predictions of the view that suggests WM information of separate subtasks may be organized as part of separate programs. Prediction d stated that if WM information that belonged to different subtasks were organized as parts of separate task related programs, then A1, A2 and A3 should have been organized as parts of different programs. This would, in turn, predict that recollection of A1 and A2 lines should not be poorer compared to the recall of A3 line, since all A lines that were kept in a given moment should be maintained separately, even though during the recollection of A1 and A2 two A lines were being maintained while only one line was being maintained during the recollection of A3. To test this prediction, we first conducted subtask A load-wise analyses on subtask A performance to compare the recollection of A1 and A2 (i.e. subtask A load of 2) to A3 recollection (i.e. subtask A load of 1). Results indicated that for subtask A accuracy, the prediction d was verified, meaning that low load condition did not differ from high load condition. Similarly, the load-wise analysis for presentation RT also did not show any difference between high and low load conditions. For subtask A recall RT, however, the high load condition was significantly higher than the low load condition. Prediction d from the separate programs view, therefore, got supported for accuracy and presentation RT.

Now, let us examine the findings of subtask A reaction time that did not seem to help to disambiguate between the two alternatives that we were examining. The first finding as such was that the subtask A recall RT was higher for the high load conditions (i.e. recollection of A1 and A2) compared to the low load condition (i.e. recollection of A3). At the first glance, this load-effect is what every WM

theory (other than program's view) would predict, including the passive store view. When we look at the episode-wise analysis for the same output measure, however, the specific prediction of the passive store view that expects a decrease in fidelity across A1 to A3 did not hold, since the recall RTs were like the following: $A1=A2$, $A1>A3$, and $A2>A3$. The effect was only limited to the number of actively maintained A lines at a given time i.e. subtask A load; high load episodes took longer to recall than low load episode. Therefore, this finding is not a direct support for the passive store view. Still, this finding is not in line with the predictions from the programs-related view, either. Let us examine this finding more thoroughly.

It is suggested in the literature that the load-effect on RT reflects that several contents are competing for selective access to the WM and to be compared to the probe [63]. When there are two A lines held in mind that are still relevant for the upcoming goals, the competition for the attentional selection might be the reason for increased recall RT in high load conditions compared to the low load condition. However, this explanation does not necessarily imply that the contents competing for the attentional selection were maintained in the same WM store. Alternatively, one might suggest that the competition was not between the individual items held (i.e. A1 and A2 during the recall of A1, and A2 and A3 during the recall of A2), but between the subtask-related programs that were actively maintained and still goal-relevant for the completion of task-episode (i.e. program for the subtask regarding mA1-rA1 and program for the subtask regarding mA2-rA2 were active during the recall of A1, and program for the subtask regarding mA2-rA2 and program for the subtask regarding mA3-rA3

were active during the recall of A2, however only one program which was the one regarding the mA3-rA3 was active by the time A3 is recalled). And once a subtask has been selectively accessed, its content was accessed without any interference from the concurrent load. This way it is possible to see a load-effect (due to the number of concurrently active task related programs) in recall RT but not in accuracy. Resolving the competition between the contents that belong to separate programs to select the probed one, even though it would slow down the processing, once it is resolved, it would not affect the accuracy [63]. Another implication of this alternative is that the attentional control of WM is mediated by the task-related programs; attentional selection to WM follows the hierarchical organization of the task-episode and subsuming program in a way that it is first performed on the level of subtasks and then within the subtasks.

The second finding that I would like to examine more thoroughly in the current experiment is that presentation RT of subtask A when examined across A1 to A3 showed that A1 was equal to A2 and A3, and A2 was higher than A3. Let me start by clarifying that this result is by no means a support for the passive store view since this view, although not directly, would at best predict that the presentation time should be $A3 > A2 > A1$. At the first glance, one might claim that the programs-related view would suggest these RTs to be the same across A1 to A3, and that is why this result does not support programs-related view, either. However, even with the possibility that all A lines were encoded into their separate WMs, there might be other dynamics that affect their memorization times. The first of such considerations might be that the hierarchical behavior suggested by the current design and the subsuming task-episode related program

got easier to execute by the time A3 was being memorized. A1 was encoded into WM, then for A2 a separate WM store was created, and A2 was encoded in it. Since a similar execution was required for A3 as A2, doing that might be easier for A3 than A2. This might indicate that once a behavior was revisited by the subsuming larger task-episode related program, which would be the largest session-related program in this current design, it was easier to perform it. Experiment 2 does not allow us to further test this idea, however, future follow up studies with variable number of steps can test whether it gets easier to perform repeated program demands within the larger task episode and whether there is a linear decrease in encoding RTs as the number of such steps are increased in a given sequence. The second alternative regarding why presentation time for A3 was the lowest might be that since by the time participants recalled A3, A1-related program was no longer goal-relevant, the content was already removed from the corresponding WM and this WM could be used to encode A3. This view implies that the larger task-episode (i.e. the session) was construed in a way that at most two WMs were created and actively maintained at a given time in Experiment 2, they are actively allocated to subtask-related programs when needed in a hierarchical manner. Although making many additional claims, both of these views that might explain the findings regarding presentation RTs across episodes are compatible with the programs-view suggesting that WM information that belongs to separate programs are maintained separately as parts of their task-related programs.

In addition, one simple consideration regarding all the RTs in Experiment 2, let it be recall RTs or presentation RTs, might be that the current design did not have

high precision in terms of capturing the reaction times due to various reasons such as it was a self-paced experiment.

Remember that the current design of Experiment 2 requires participants to actively organize A lines in their WM to recollect not the last A line but the one before that. This active organization is the key to keep the A line to be recalled from the interference from the last A line presented. Active organization of information is a feature of WM that is well-recognized [12]. Findings of experiment 2, as well, suggests that A lines did not interfere with the A lines presented prior to it, they were actively organized to be retrieved at the junctions that are tied to different goals. In addition to active organization of WM information, current design also required an active removal [13] of no longer goal relevant A lines from the WM after they are recalled. If this active removal would not happen, we would expect to see performance declining across four episodes of a session due to piling up of A lines memorized. Instead, we saw that the subtask A performance did not show a decrease throughout the episodes. Active organization of WM information and removal of no longer goal relevant information were enabling features of WM that helped to drive the findings in Experiment 2.

When all the findings of Experiment 2 were taken into consideration along with the discussion above, evidence was largely in favor of programs-related view, and none of the predictions from passive store view held. In conclusion, collective evidence from Experiment 1 and Experiment 2 of the current thesis suggests that WM information that belongs to separate task programs might be

maintained in their corresponding WMs without interfering with the information that belongs to other task programs. And, this, in turn, might be another way of circumventing WMC limits beyond what is suggested by the literature, just like chunking.

In Experiment 3 and Experiment 4, we investigated some factors that might affect the maintenance and processing of WM information that belongs to different task-related programs.

Chapter 4

Experiment 3

There are different phases to a short-term or WM system: encoding, maintenance, and retrieval. The crucial implication of collective evidence from Experiment 1 and Experiment 2 was that WM information of different subtasks were maintained as parts of their task-related programs, separately. This implies that information belonging to different subtasks were encoded into their corresponding WMs separately. The details regarding this encoding, however, remains unclear.

Consolidation of sensory information into stable memory traces is an important attentional-bottleneck to the memory encoding [82]. Consolidation of WM information, as mentioned in the general introduction of this thesis, is the attention-based process where the temporary sensory representations are transformed into stable memory traces or representations that can be further processed, specifically manipulated, and retrieved even after some delay period [53]. Consolidation is usually studied concerning the benefit of free time after presenting a set of stimuli [54]. Such a retention interval is thought to reflect the time course for the consolidation. Nonetheless, the attentional control of WM research suggests that consolidation of the WM items into their stable memory traces or representations is a time-dependent process, and it is implied that devoting more time to actively maintaining the WM traces that are newly formed

would lead to a better encoding of these traces into long term episodic memory mainly because it would strengthen the activation of memory traces [83].

Decreasing the memory encoding time by simply applying a time constraint on the stimulus onset asynchrony (i.e. the time between the onset of stimulus presentation and the onset of another stimulus) is found to decrease the memory performance. In a visual WM study of Knöchel et al. [84], participants were presented with an array of colored squares for 100 ms, given a variable stimulus onset asynchrony time, then immediately presented with a pattern mask, and probed with a new array to detect whether it was a match to the target array or not. Crucially, the stimulus onset asynchrony and the number of squares in the memory array were manipulated orthogonally. They showed that both the memory load and stimulus onset asynchrony affected memory performance, and stimulus onset asynchrony lost its effect on performance somewhere between 400 to 800 milliseconds.

The time that is required for consolidation was found to be varying from study to study. For instance, for single objects even as low as 50 ms was suggested [85]. For an array of objects that consisted of three to five objects in the study of Knöchel et al., for example, it was somewhere between 400 to 800 milliseconds [84]. Indeed, it is now well-known that time and cognitive effort required for consolidation might be higher for higher loads of information.

In line with these suggestions, decreasing the time that WM information is actively maintained in the attentional focus right after it is presented should lead

to inadequate consolidation of memory traces. And, if attention-based consolidation is part of the mechanism that encodes the WM information to its corresponding task-related program, disruption in consolidation processes should affect the interference-free maintenance of separate WM information in separate programs.

In Experiment 3, we wanted to investigate this possibility. To test the effect of limited time for memory consolidation on the lack of across-subtask load-effect that we have observed in Experiment 1 and Experiment 2, we modified the experimental paradigm of Experiment 1 to include a time constraint for the presentation of stimulus. Initially, participants kept one or two A lines in their mind (mA). Then they were presented with three B lines (mB), and probed to recall the orientation of one of them (rB). Finally, they were probed to recall the orientation of one of the A lines they saw (rA). Most crucially, they were presented with the stimuli in mA and mB for a brief amount of time which was decided to be 700 ms. This design allowed us to see whether subtask A load interferes with subtask B performance when the self-paced nature of the study is abandoned and the stimulus encoding time is fixed to be a brief amount of time.

4.1 Methods

4.1.1 Procedure

Experimental procedure was similar to the Experiment 1, except that the reimbursement was done as either course credit or a payment of \$22-30 based on the performance on the task.

4.1.2 Participants

42 healthy participants (22 females) within an age group of 20-26 ($M=21.7$, $SD=2.14$) were recruited. Participants are given digital copies of the informed consent and they sign them with their e-signature before participation in the experiment. Participants were reimbursed with course credits or with payment of ₺22-30 based on their performance.

4.1.3 Experimental Paradigm

The experimental paradigm of Experiment 3 was similar to Experiment 1 except the following: In experiment 3, participants had only 700 ms to memorize the presented lines during mA and mB, so the experiment was not completely self-paced (Figure 4.1). Also, in Experiment 3, subtask B WM load was not manipulated, and during all mBs, 3 lines were presented. This left us with two conditions for subtask A and subtask B WM loads for a given episode: [1,3] and [2,3]. These conditions are pseudorandomized in occurrence across participants and across episodes of a session. Most crucially, a difference between these conditions [i.e. effect of subtask A WM load on subtask B performance] would indicate across-subtasks load effect in Experiment 3.

Another change in the design was that participants were given monetary incentives based on their deviation in each recall phase, in each recall, they received a large compensation (in kuruş) if their degree of deviation is was between 0 and 45 degrees, with the compensation decreasing as it got closer to the 0 degree. If their deviation was greater than 45 degrees, they were not compensated for that recall. They got feedback regarding the monetary value of

their immediate recall, in addition to the feedback regarding the deviation from the target in degrees.

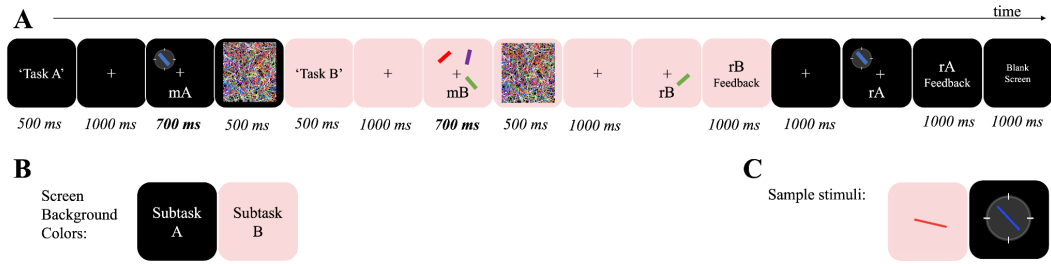


Figure 4.1. Experimental Paradigm of Experiment 3. A. The order of events and screens in Experiment 3. B. Background colors were different for subtasks. C. Illustrates some sample lines in different backgrounds during the memorization phase. Please, note that the subtask A line was presented with a circle around it, while the subtask B line was not.

4.1.4 Data Analysis

All analyses were conducted on MATLAB R2021b, and figures illustrating analysis results were created in Prism GraphPad Software [version 9.4.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com] by using the mean and 95% Confidence Interval [CI] information gathered from MATLAB analyses. 95% CIs were calculated by using Cousineau-Morey method for within-participant or repeated measures designs [77; 78]. Descriptive statistics tables and paired-sample t-test tables are gathered from Jamovi [79; 80].

In this experiment, as mentioned earlier, the number of lines presented in a subtask B mB is not manipulated, only the number of A lines to keep was varying from episode to episode. Hence, the emphasis was on the effect of withheld subtask A WM load on subtask B WM task performance (i.e. across-subtasks load

effect), as well as, the effect of subtask A load on subtask A performance (i.e. within-subtask load effect). This led us to have two different levels for our independent variable, namely subtask A WM load; load of 1 (low) and load of 2 (high). Data from these two groups are compared in paired samples T-test analyses. Since the presentation time of subtask A and subtask B lines were fixed, there was no output measure regarding presentation time in this experiment.

4.2 Results

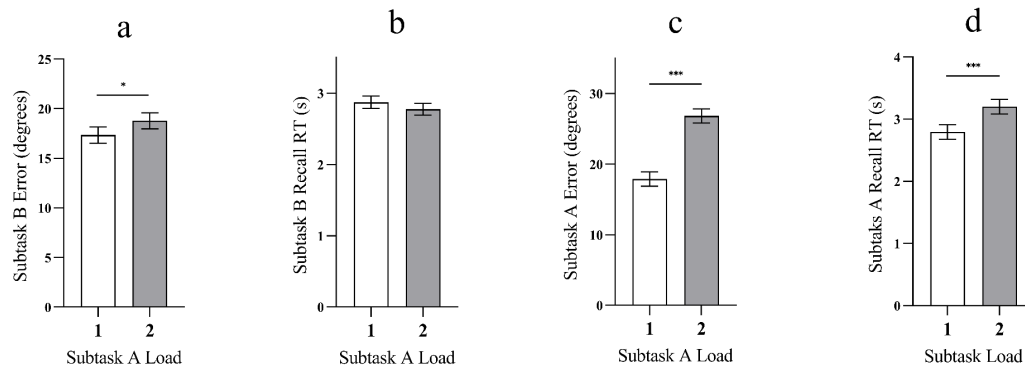


Figure 4.2. Paired sample T-test results comparing subtask A WM load of 1 and of 2 in Experiment 3. Output measures are: (a) Subtask B Error (degrees), (a) Subtask B Recall RT (s), (c) Subtask A Error (degrees), (d) Subtask A Recall RT [s].

4.2.1 Subtask B Error

Subtask B error was calculated as the mean deviation of the orientation recalled in rB from the target line's orientation presented in corresponding mB, in degrees. Table 4.1 shows the descriptives for subtask A load conditions. Paired samples t-test analysis (Table 4.2) revealed that subtask B error was significantly higher for the rBs where participants needed to withhold 2 A lines ($M = 18.6$, $SD = 5.92$)

compared to 1 A line ($M = 17.2$, $SD = 5.46$) ($t(41) = -2.55$, $p = 0.02$, $CI\ 95\% = [-2.50, -0.29]$, $d = -0.40$). For higher subtask A load condition, subtask B error was higher. Figure 4.2.a illustrates the results.

Table 4.1. Descriptives for Subtask B Error [degrees] in Experiment 3

	Subtask A Load: 1 [low]	Subtask B Load: 2 [high]
N	42	42
Missing	0	0
Mean	17.2	18.6
Std. error mean	0.842	0.913
95% CI mean lower bound	15.5	16.8
95% CI mean upper bound	18.8	20.4
Median	16.4	17.1
Mode	8.20 ^a	11.5 ^a
Standard deviation	5.46	5.92
Minimum	8.20	11.5
Maximum	33.7	35.5
Skewness	0.710	1.08
Std. error skewness	0.365	0.365

^a More than one mode exists, only the first is reported

Table 4.2. Paired Samples T-test for Subtask B Error [degrees] in Experiment 3

		statistic	df	p	Mean difference	SE difference	95% Confidence Interval		Effect Size	95% Confidence Interval			
							Lower	Upper		Lower	Upper		
AL: 1 [low]	AL: 2 [high]	Student's t	-2.55	41.0	0.015	-1.40	0.548	-2.50	-0.289	Cohen's d	-0.393	-0.705	-0.0767

4.2.2 Subtask B Recall RT

Subtask B recall RT was calculated as the mean of the amount of time participants took in rB screen, in seconds. Table 4.3 shows the descriptives for subtask A load conditions. Paired samples t-test analysis (Table 4.4) comparing withheld A load of 1 ($M = 2.87$, $SD = 0.916$) and withheld A load of 2 ($M = 2.77$, $SD = 0.809$) on subtask B recall RT revealed that two means were not significantly different ($t(41)$

= 1.67, $p = 0.10$, $CI\ 95\% = [-0.02, 0.22]$, $d = 0.26$). Figure 4.2.b illustrates the results.

Table 4.3. Descriptives for Subtask B Recall RT [s] in Experiment 3

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	42	42
Missing	0	0
Mean	2.87	2.77
Std. error mean	0.141	0.125
95% CI mean lower bound	2.60	2.53
95% CI mean upper bound	3.15	3.02
Median	2.61	2.51
Mode	1.86 ^a	1.70 ^a
Standard deviation	0.916	0.809
Minimum	1.86	1.70
Maximum	6.36	4.85
Skewness	1.84	1.06
Std. error skewness	0.365	0.365

^a More than one mode exists, only the first is reported

Table 4.4. Paired Samples T-test for Subtask B Error [degrees] in Experiment 3

										95% Confidence Interval		95% Confidence Interval		
										Lower	Upper	Effect Size	Lower	Upper
AL: 1 [low]	AL: 2 [high]	Student's t	statistic	df	p	Mean difference	SE difference							
			1.67	41.0	0.103	0.0989	0.0593	-0.0209	0.219	Cohen's d	0.257	-0.0517	0.563	

4.2.3 Subtask A Error

Subtask A error was calculated as the mean deviation of the orientation recalled in rA from the target line's orientation presented in corresponding mA, in degrees. Table 4.5 shows the descriptives for subtask A load conditions. Paired samples t-test analysis (Table 4.6) comparing subtask A load of 1 ($M = 17.5$, $SD = 9.36$) and subtask A load of 2 ($M = 26.6$, $SD = 8.27$) on subtask A error revealed that error was significantly higher in high load condition compared to the low ($t(41) =$

-9.43, $p < 0.001$, $CI\ 95\% = [-11.03, -7.14]$, $d = -1.46$). Figure 4.2.c illustrates the results.

Table 4.5. Descriptives for Subtask A Error [degrees] in Experiment 3

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	42	42
Missing	0	0
Mean	17.5	26.6
Std. error mean	1.44	1.28
95% CI mean lower bound	14.7	24.1
95% CI mean upper bound	20.4	29.1
Median	15.6	26.5
Mode	6.45	12.9 ^a
Standard deviation	9.36	8.27
Minimum	6.45	12.9
Maximum	47.5	54.3
Skewness	1.17	0.714
Std. error skewness	0.365	0.365

^a More than one mode exists, only the first is reported

Table 4.6. Paired Samples T-test for Subtask A Error [degrees] in Experiment 3

		statistic	df	p	Mean difference	SE difference	95% Confidence Interval		Effect Size	95% Confidence Interval			
							Lower	Upper		Lower	Upper		
AL: 1 [low]	AL: 2 [high]	Student's t	-9.43	41.0	<.001	-9.09	0.963	-11.0	-7.14	Cohen's d	-1.46	-1.89	-1.02

4.2.4 Subtask A Recall RT

Subtask A recall RT was calculated as the mean of the amount of time participants took in rA screen, in seconds. Table 4.7 shows the descriptives for subtask A load conditions. Paired samples t-test analysis (Table 4.8) comparing subtask A load of 1 ($M = 17.5$, $SD = 9.36$) and subtask A load of 2 ($M = 17.5$, $SD = 9.36$) on subtask A recall RT revealed that in high subtask A load condition, subtask B took longer to recall compared to the low subtask A load condition ($t(41) = -4.93$, $p < 0.001$, $CI\ 95\% = [-0.57, -0.24]$, $d = -0.76$). Figure 4.2.d illustrates the results.

Table 4.7. Descriptives for Subtask A Recall RT [s] in Experiment 3

	Subtask A Load: 1 [low]	Subtask A Load: 2 [high]
N	42	42
Missing	0	0
Mean	2.79	3.20
Std. error mean	0.130	0.170
95% CI mean lower bound	2.54	2.87
95% CI mean upper bound	3.05	3.54
Median	2.63	3.02
Mode	1.58 ^a	1.41 ^a
Standard deviation	0.839	1.10
Minimum	1.58	1.41
Maximum	5.00	6.20
Skewness	1.03	0.901
Std. error skewness	0.365	0.365

^a More than one mode exists, only the first is reported

Table 4.8. Paired Samples T-test for Subtask A Recall RT [s] in Experiment 3

		statistic	df	p	Mean difference	SE difference	95% Confidence Interval		Effect Size	95% Confidence Interval		
							Lower	Upper		Lower	Upper	
WL: 1 [low]	WL: 2 [high]	Student's t	-4.93	41.0	<.001	-0.408	0.0828	-0.576	-0.241	0.761	-1.103	-0.413

4.3 Discussion

In Experiment 3, the number of lines withheld in relation to subtask A affected the subtask A performance (i.e. within-subtask load effect) and crucially subtask B accuracy, as well, (i.e. across-subtask load effect) indicating that in Experiment 3, interference across WMCs of different subtasks was present.

When the self-paced nature of Experiment 1 is abandoned, it seems the suggestion that separate task-related programs maintain their WM information in their corresponding WMs no longer holds. When combined with Experiment 1 and Experiment 2, this finding suggests that consolidation of a subtask's

corresponding WM information into an entity related to its task-related program requires time under attentional execution, and it is not an automatic process.

One possibility is that simultaneously presented stimuli that belong to a subtask are consolidated together to create a single entity that is closely tied to its corresponding task program, and they might be maintained in it separately from the other subtask's WM information. Indeed, similarly, previous research suggested that the functional meaning behind the consolidation might be that it simplifies the relationships between the items to be stabilized by identifying their familiar aspects and helping create chunked information [54]. It is also suggested that such a process provides the memory traces with an episodic distinctiveness [86] by binding semantic and perceptual features into a token representation and providing the endpoints of temporal episodes [87]. Also, a function of binding content to its context is suggested for the initial attentional processes over a set of transient sensory inputs [81]. In addition to these suggestions, we suggest that another functional role of attention-based consolidation might be that adequate consolidation of WM traces might be a part of the mechanism by which WM information is encoded into their task-related program.

We suggested that WM information's being organized as part of separate task related programs might be another way of circumventing WMC limits, just like chunking. Chunking literature suggests that by grouping the information together (i.e. creating chunks) more information than the limits suggested by the literature can be kept [30]. Chunking is seen as an attentional time-based mechanism that enhances memory performance [30]. On the other hand, a study that investigated

the relationship between the chunking and WM showed that in a complex span WM task, chunking competes with other attention-based mechanisms especially during the encoding phase of WM information [88]. This finding also suggested that chunking links WM to LTM precisely at the encoding of information, and does so in an attentional time-based manner [88]. Just like chunking, encoding WM information to corresponding subtask's program might be an attentional time-based mechanism, and might lead to memory enhancement by circumventing WMC limits in a new way.

Chapter 5

Experiment 4

Experiment 1 and Experiment 2 collectively suggested that when identical WM tasks are organized in a hierarchical manner, the WM information that belongs to different task related programs might be maintained in their corresponding WMs separately. And, evidence from Experiment 3 suggested that the encoding of information to its corresponding subtask related program and WM is a time-dependent and deliberate process, and not automatic. However, on the other hand, retrieval of WM information from those separate WMs accomplished in an interference-free manner to cause the lack of across-subtask load effect is still unclear.

In many formal models of WM, retrieval in WM is described as a competition between the items that are held concurrently [89]. Some models argue that the items are held active from the presentation to retrieval [90], whereas, others assume that the WM representations are reactivated during the retrieval phase by the contextual cues that cue the target item [91]. Contextual cues that are referred to here can be present task context (i.e. cues discriminating subtask A from subtask B), or the spatial position of the items (i.e. cues discriminating oriented lines within a subtask set) in our experimental paradigms. It is suggested that when the contextual cues are not distinctive enough to selectively cue the retrieval

of the target WM representation, other retrieval candidates are activated as strongly as the target, and ‘interference by confusion’ occurs [37]. Similarly, when the WM items are distributed representations (when they have more than one features that relates to them, and they are maintained as a vector of those features, i.e. oriented lines in this thesis’ experimental paradigms coded in relation the their orientation, spatial position, and color) the shared features between the WM representations could be overwritten in one of them [37]. All in all, the interference hypothesis of WMC suggests that capacity limitations arise from the interference between the WM representations, and interference accounts that I mentioned places a special emphasis on the similarity of individual items or contextual cues as the driving factors for interference in WM. This implies that increasing similarity between the items in a given subtask set or the contextual cues that discriminate subtask A from subtask B should lead to increased interference.

In Experiment 4, we wanted to investigate this possibility in our newly introduced extended task paradigm. Note that in Experiment 1, although the subtask-related program A was almost identical to subtask-related program B, the presentation mode of stimuli that belonged to different subtasks differed in a way that subtask A lines were presented within a circle while the subtask B lines were presented freely. This difference arguably made the stimuli that belonged to different subtasks more distinct, and led to concurrent subtask related programs identities’ being even more distinguishable. In Experiment 4, we wanted to modify Experiment 1 in a way that the subtask A and subtask B demands are even more similar to each other by using the same presentation mode in both subtasks.

The aim was to explore the effect of increased similarity between subtask demands on the previous finding that WMCs did not interfere across subtasks.

5.1 Methods

5.1.1 Procedure

Experimental procedure was similar to Experiment 1 with some exceptions. Instead of browsing the experiment link through their local computer, participants were provided with an executable program file (with the extension of .exe) of the experiment that they can run on their local computer, and the data is collected by this program, and saved directly to the local computer, and since such executable files are not compatible with some operating systems, participants were selected amongst the pool if they had access to a computer operating on Microsoft Windows. When they are done with the experiment, they have sent the data file back to the experimenter via email. They were not able to see the source code, and could not manipulate the data file manually. Another exception is that the reimbursement was done as either course credit or a payment of ₦20 regardless of the performance on the task.

5.1.2 Participants

32 healthy participants (20 females) were recruited. 4 of them (1 female) were eliminated from further processes due to various reasons [not completing the whole experiment (N=1), the data collected in the local computer were disrupted by an unknown reason (N=3)] leaving the final sample as 28 participants (19 females) within an age group of 18-25 ($M=22$, $SD=1.86$). Participants were given digital copies of the informed consent and they signed them with their e-signature prior to

the participation. They were reimbursed with course credits or with a payment of £20.

5.1.3 Experimental Paradigm

Experiment 4 was similar to Experiment 1: In a given episode, participants were presented with 1 or 2 A lines to keep in mind (mA), then they were presented with 2 or 3 B lines to memorize (mB), and are probed to recall one of the B lines presented (rB). Finally, they were probed to recall an A line's orientation (rA) (Figure 5.1.A). One deviance from the Experiment 1 was that subtask A lines were presented without a circle around them, just like subtask B lines, increasing the similarity of withheld WM content and concurrent WM content (Figure 5.1.C). During the recall phases (rA and rB), they were presented with a circle around them, to cue that the screen was a recall screen. Just like in Experiment 1, there were four conditions for pairs of WM loads of subtask A and subtask B loads in an episodes ([1,2], [1,3], [2,2], [2,3]), load condition pairs occurred equally frequently for a given participant, and their order of occurrence is pseudo-randomized across participants.

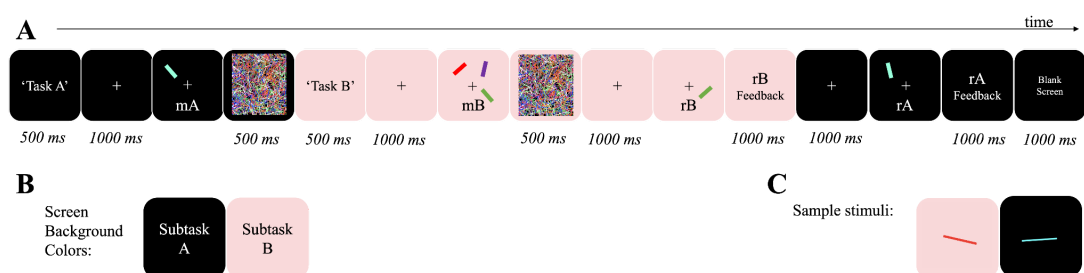


Figure 5.1. Experimental paradigm of Experiment 4. A. The order of events and screens in Experiment 4. B. Screen background colors were different for subtask A and subtask B related parts. Subtask A was executed in a pink screen whereas subtask B was executed in a black screen. C. Illustration of some sample lines in different backgrounds during the memorization phase. Please note that the

presentation mode is the same for both of the tasks (neither of them is presented with a circle around them, unlike previous experiments).

5.1.4 Data Analysis

Raw data is prepared for statistical analyses on MATLAB R2021b, repeated measures ANOVA and post-hoc comparisons are performed on MATLAB, as well. Figures illustrating analysis results were created in Prism GraphPad Software [version 9.4.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com] by using the mean and 95% Confidence Interval (CI) information gathered from MATLAB analyses. 95% CIs were calculated by using Cousineau-Morey method for within-participant or repeated measures designs [77; 78]. Descriptive statistics tables and ANOVA tables are retrieved from Jamovi [79; 80].

Considering the similarity between Experiment 1 and current experiment's structure, all the data analysis logic was followed from Experiment 1. We conducted a 2x2 Repeated Measures Analysis of Variance (ANOVA) to see the effect of Subtask A Load and Subtask B Load on all output measures of interest: subtask b error, subtask b recall reaction time, subtask b presentation time, subtask a error, subtask a recall reaction time, and subtask a presentation time.

In Experiment 4, too, we had four pairs of load conditions, and we pre-determined a subset of pairwise comparisons we were interested in for each output measure, to perform planned pairwise comparisons for only this subset [with Bonferroni correction], to increase statistical power by decreasing the number of

pairwise comparisons performed. This subset consisted of the following contrasts: ([1,2] vs [1,3]), ([2,2] vs [2,3]), ([1,2] vs [2,2]), and ([1,3] vs [2,3]). Again, the reason for these comparisons being selected was to investigate the effect of manipulating the number of lines that are held in mind for a certain task on the performance of the other task when the latter's WM load was kept constant.

5.2 Results

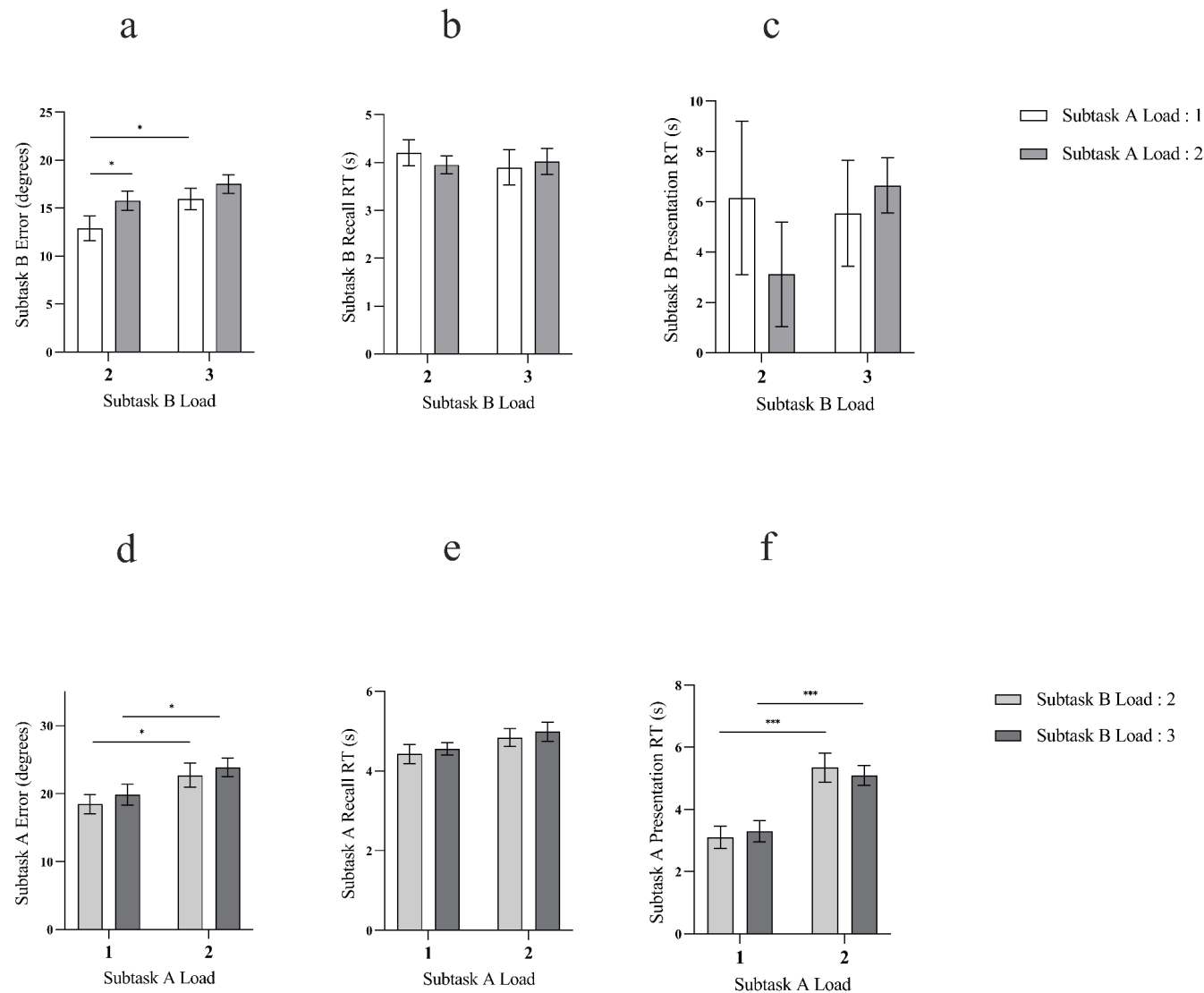


Figure 5.2. Illustration of the 2x2 repeated measures ANOVA and planned contrast with Bonferroni correction results for Experiment 4, for the output measures of subtask B error (a), subtask B recall RT (b), subtask B presentation RT (c), subtask A error (d), subtask A recall (e), subtask A presentation RT (f). Error bars: 95% CI. The significance stars are only present (or absent) for the pairwise comparisons of interest.

5.2.1 Subtask B Error

The subtask B error was calculated as the mean deviation of the recall in rB from the target shown in the respective mB, in degrees. Table 5.1 summarizes the descriptive statistics for subtask B error. The repeated measures ANOVA revealed that (Table 5.2) both of the main effects of subtask B load (Subtask B WM Load: $F(1,27) = 9.22, p = 0.005, \eta_p^2 = 0.255$) and subtask A load (Subtask A WM Load: $F(1,27) = 9.30, p = 0.005, \eta_p^2 = 0.255$) were present on concurrent task error. The interaction effect of subtask A and subtask B load was not significant either (Interaction: Subtask A Load x Subtask B Load: $F(1,27) = 1.51, p = 0.23, \eta_p^2 = 0.053$). Both across-subtask and within-subtask load-effect or interference was observed on subtask B error.

Pairwise comparisons with Bonferroni correction (Table 5.3) revealed that the groups [1,2] vs [1,3] (Mean Difference= -3.06, SE= 1.09, $t(44) = -2.813, p=0.036$), and [1,2] vs [2,2] (Mean Difference: -2.88, SE: 0.99, $t(44) = -2.904, p=0.029$) significantly differed. Comparisons of the means for the groups [2,2] vs [2,3] (Mean Difference: -1.74, SE: 0.99, $t(27) = -2.172, p=0.155$) and [1,3] vs [2,3] (Mean Difference: -1.56, SE: 0.80, $t(27) = -1.930, p=0.256$) did not yield significant results. The results revealed that not only increasing subtask B Load but also increasing the subtask A Load increased subtask B error. Figure 5.2.a illustrates these results.

Table 5.1. Descriptives for Subtask B Error [degrees] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	12.9	16.0	15.8	17.5

Table 5.1. Descriptives for Subtask B Error [degrees] in Experiment 4

	1_2	1_3	2_2	2_3
Std. error mean	1.78	1.72	1.98	1.90
95% CI mean lower bound	9.40	12.6	11.9	13.8
95% CI mean upper bound	16.4	19.3	19.7	21.2
Median	9.31	13.4	12.7	14.8
Mode	7.48 ^a	4.57 ^a	3.90 ^a	10.9
Standard deviation	9.44	9.08	10.5	10.0
Minimum	3.19	4.57	3.90	3.24
Maximum	49.2	39.5	41.3	39.5
Skewness	2.42	0.947	1.35	1.09
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.2. Within Subjects Effects for Subtask B Error [degrees] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	138.3	1	138.34	9.30	0.005	0.013	0.013	0.256
Residual	401.5	27	14.87					
Subtask B Load	161.9	1	161.92	9.22	0.005	0.015	0.015	0.255
Residual	474.1	27	17.56					
Subtask A Load * Subtask B Load	12.3	1	12.25	1.51	0.229	0.001	0.001	0.053
Residual	218.7	27	8.10					

Note. Type 3 Sums of Squares

Table 5.3. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Error [degrees] in Experiment 4

Comparison					Mean Difference	SE	df	t	p_{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-3.066	1.090	27.0	-2.813	0.036
		-	2 [high]	2 [low]	-2.884	0.993	27.0	-2.904	0.029
	3 [high]	-	2 [high]	3 [high]	-1.561	0.809	27.0	-1.930	0.256
2 [high]	2 [low]	-	2 [high]	3 [high]	-1.743	0.803	27.0	-2.172	0.155

5.2.2 Subtask B Recall RT

The subtask B recall RT was calculated as the time from the probe is shown and button response is made in rB, in seconds. Table 5.4 summarizes the descriptive

statistics for subtask B recall RT. 2x2 Repeated Measures ANOVA (Table 5.5) revealed that none of the effects were significant (Subtask B Load: $F(1,27) = 1.39$, $p = 0.25$, $\eta^2_p = 0.049$) (Subtask A Load: $F(1,27) = 0.35$, $p = 0.56$, $\eta^2_p = 0.013$) (Interaction: Subtask A Load x Subtask B Load: $F(1,27) = 0.50$, $p = 0.49$, $\eta^2_p = 0.018$). None of the planned pairwise comparisons yielded significant results (Table 5.6). The results are illustrated in Figure 5.2.b.

Table 5.4. Descriptives for Subtask B Recall RT [s] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	4.21	3.90	3.95	4.03
Std. error mean	0.521	0.371	0.422	0.498
95% CI mean lower bound	3.19	3.17	3.12	3.05
95% CI mean upper bound	5.23	4.62	4.78	5.00
Median	3.37	3.68	3.39	3.38
Mode	1.32 ^a	1.47 ^a	0.970 ^a	1.17 ^a
Standard deviation	2.76	1.96	2.24	2.64
Minimum	1.32	1.47	0.970	1.17
Maximum	13.0	11.9	12.6	13.2
Skewness	2.38	2.81	2.45	2.42
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.5. Within Subjects Effects for Subtask B Recall RT [s] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	0.113	1	0.113	0.346	0.561	0.000	0.000	0.013
Residual	8.833	27	0.327					
Subtask B Load	0.389	1	0.389	1.390	0.249	0.001	0.001	0.049
Residual	7.555	27	0.280					
Subtask A Load * Subtask B Load	1.026	1	1.026	0.499	0.486	0.002	0.002	0.018
Residual	55.576	27	2.058					

Note. Type 3 Sums of Squares

Table 5.6. Post Hoc Comparisons -Subtask A Load * Subtask B Load - for Subtask B Recall RT [s] in Experiment 4

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	0.3093	0.336	27.0	0.920	1.000
		-	2 [high]	2 [low]	0.2550	0.235	27.0	1.087	1.000
	3 [high]	-	2 [high]	3 [high]	-0.1278	0.340	27.0	-0.377	1.000
2 [high]	2 [low]	-	2 [high]	3 [high]	-0.0736	0.232	27.0	-0.316	1.000

5.2.3 Subtask B Presentation RT

The subtask B presentation RT was calculated as the mean amount of time participants took to memorize the lines shown in the screen mB, [the time between the mB started and participant made the button response to move onto the next screen], in seconds. Table 5.7 summarizes the descriptive statistics for subtask B presentation RT. 2x2 Repeated Measures ANOVA (Table 5.8) revealed that none of the effects were significant (Subtask B Load: $F(1,27) = 2.60$, $p = 0.12$, $\eta^2_p = 0.088$) (Subtask A Load: $F(1,27) = 2.67$, $p = 0.114$, $\eta^2_p = 0.09$) (Interaction: Subtask A Load x Subtask B Load: $F(1,27) = 0.94$, $p = 0.34$, $\eta^2_p = 0.034$). None of the planned pairwise comparisons yielded significant results (Table 5.9). The results are illustrated in Figure 5.2.c.

Table 5.7. Descriptives for Subtask B Presentation RT [s] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	6.15	5.54	3.12	6.65
Std. error mean	2.67	0.941	0.336	1.74
95% CI mean lower bound	0.910	3.70	2.46	3.25
95% CI mean upper bound	11.4	7.38	3.78	10.1
Median	2.88	3.51	2.68	3.08
Mode	0.730 ^a	0.701 ^a	0.641 ^a	1.10 ^a
Standard deviation	14.1	4.98	1.78	9.20
Minimum	0.730	0.701	0.641	1.10
Maximum	77.4	25.1	7.14	46.8
Skewness	5.09	2.52	1.05	3.56

Table 5.7. Descriptives for Subtask B Presentation RT [s] in Experiment 4

	<u>1_2</u>	<u>1_3</u>	<u>2_2</u>	<u>2_3</u>
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.8. Within Subjects Effects for Subtask B Presentation RT [s] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	25.7	1	25.67	2.669	0.114	0.003	0.003	0.090
Residual	259.7	27	9.62					
Subtask B Load	59.9	1	59.86	2.599	0.119	0.007	0.007	0.088
Residual	621.7	27	23.03					
Subtask A Load * Subtask B Load	119.9	1	119.93	0.943	0.340	0.014	0.014	0.034
Residual	3432.9	27	127.15					

Note. Type 3 Sums of Squares

Table 5.9. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask B Presentation RT [s] in Experiment 4

Comparison					Mean Difference	SE	df	t	p _{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	0.608	2.791	27.0	0.218	1.000
		-	2 [high]	2 [low]	3.027	2.693	27.0	1.124	1.000
	3 [high]	-	2 [high]	3 [high]	-1.112	1.587	27.0	-0.701	1.000
2 [high]	2 [low]	-	2 [high]	3 [high]	-3.532	1.713	27.0	-2.062	0.196

5.2.4 Subtask A Error

The subtask A error was calculated as the mean deviation of the recall in rA from the target shown in respective mA, in degrees. Table 5.10 summarizes the descriptive statistics for subtask A error. The repeated measures ANOVA revealed that (Table 5.11) main effect of subtask A load on subtask A error was significant (Subtask A Load: $F(1,27)=17.85$, $p<0.001$, $\eta^2_p = 0.398$) whereas the main effect of subtask B load was not significant (Subtask B Load: $F(1,27)=1.65$, $p=0.21$, $\eta^2_p =$

0.057). The interaction effect was not significant either (Interaction: Subtask A Load x Subtask B Load: $F(1,27)=0.02$, $p=0.89$, $\eta^2_p = 0.001$).

Pairwise comparisons with Bonferroni correction (Table 5.12) revealed that the groups showed a significant difference when the concurrent WM task load was kept constant, for comparisons [1,2] vs [2,2] (Mean Difference = -4.28, SE = 1.46, $t(27) = -2.937$, $p = 0.027$), and [1,3] vs [2,3] (Mean Difference = -4.02, SE = 1.25, $t(27) = -3.223$, $p=0.013$). When the Subtask A Load has kept constant the comparisons of [1,2] vs [1,3] (Mean Difference = -1.40, SE = 1.29, $t(27) = -1.086$, $p \cong 1.00$), and [2,2] vs [2,3] did not yield significant results (Mean Difference = -1.13, SE = 1.43, $t(27) = -0.795$, $p \cong 1.00$). Figure 5.2.d illustrates these results. Results showed that only changing the number of subtask A lines affected subtask A error.

Table 5.10. Descriptives for Subtask A Error [degrees] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	18.4	19.8	22.7	23.9
Std. error mean	2.10	2.50	2.64	2.10
95% CI mean lower bound	14.3	14.9	17.5	19.7
95% CI mean upper bound	22.6	24.7	27.9	28.0
Median	14.8	17.4	19.6	22.6
Mode	4.52 ^a	6.81	3.29 ^a	7.24 ^a
Standard deviation	11.1	13.2	14.0	11.1
Minimum	4.52	4.29	3.29	7.24
Maximum	43.3	53.0	55.1	49.0
Skewness	0.993	1.01	0.930	0.615
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.11. Within Subjects Effects for Subtask A Error [degrees] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	481.361	1	481.361	17.8462	< .001	0.028	0.028	0.398
Residual	728.263	27	26.973					
Subtask B Load	44.949	1	44.949	1.6450	0.211	0.003	0.003	0.057
Residual	737.768	27	27.325					
Subtask A Load * Subtask B Load	0.480	1	0.480	0.0196	0.890	0.000	0.000	0.001
Residual	660.060	27	24.447					

Note. Type 3 Sums of Squares

Table 5.12. Post Hoc Comparisons -Subtask A Load * Subtask B Load - for Subtask A Error [degrees] in Experiment 4

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A Load	Subtask B Load	Subtask A Load	Subtask B Load						
1 [low]	2 [low]	- 1 [low]	3 [high]		-1.40	1.29	27.0	-1.086	1.000
		- 2 [high]	2 [low]		-4.28	1.46	27.0	-2.937	0.027
	3 [high]	- 2 [high]	3 [high]		-4.02	1.25	27.0	-3.223	0.013
2 [high]	2 [low]	- 2 [high]	3 [high]		-1.14	1.43	27.0	-0.795	1.000

5.2.5 Subtask A Recall RT

The subtask A recall reaction time was calculated as the time from the probe is shown and button response is made in the screen rA, in seconds. Table 5.13 summarizes the descriptive statistics for subtask A recall RT. 2x2 Repeated Measures ANOVA (Table 5.14) showed that similar to the withheld task error, only the main effect of subtask A load was present (Subtask A Load: $F(1,27)=7.84$, $p=0.009$, $\eta^2_p = 0.225$) whereas the main effect of subtask B load (Subtask B Load: $F(1,27)=1.16$, $p=0.29$, $\eta^2_p = 0.041$) and the interaction effect (Interaction: Subtask B Load x Subtask A Load: $F(1,27)=0.009$, $p=0.92$, $\eta^2_p = 0.000$) were not significant.

Pairwise comparisons with Bonferroni correction (Table 5.15) revealed that all the pairwise comparisons of interest yielded non-significant results ([1,3] vs [2,3]:

Mean Difference = -0.43, SE = 0.17, $t(27) = -2.43$, $p=0.088$), ([1,2] vs [1,3]:
Mean Difference = -0.13, SE = 1.16, $t(27) = -0.787$, $p \cong 1.00$), ([1,2] vs [2,2]:
Mean Difference = -0.41, SE = 0.22, $t(27) = -1.872$, $p = 0.288$), ([2,2] vs [2,3]:
Mean Difference = -0.15, SE = 0.20, $t(27) = -0.739$, $p \cong 1.00$). Figure 5.2.e
illustrates these results.

Table 5.13. Descriptives for Subtask A Recall RT [s] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	4.43	4.55	4.84	4.99
Std. error mean	0.409	0.488	0.505	0.553
95% CI mean lower bound	3.63	3.60	3.85	3.90
95% CI mean upper bound	5.23	5.51	5.83	6.07
Median	3.80	4.00	4.11	4.12
Mode	1.81 ^a	1.18 ^a	0.889 ^a	1.07 ^a
Standard deviation	2.16	2.58	2.67	2.93
Minimum	1.81	1.18	0.889	1.07
Maximum	11.0	12.9	13.3	14.3
Skewness	1.39	1.72	1.47	1.76
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.14. Within Subjects Effects for SubtaskA Recall RT [s] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	4.96872	1	4.96872	7.84216	0.009	0.007	0.007	0.225
Residual	17.10695	27	0.63359					
Subtask B Load	0.53572	1	0.53572	1.16121	0.291	0.001	0.001	0.041
Residual	12.45634	27	0.46135					
Subtask A Load * Subtask B Load	0.00442	1	0.00442	0.00921	0.924	0.000	0.000	0.000
Residual	12.94557	27	0.47947					

Note. Type 3 Sums of Squares

Table 5.15. Post Hoc Comparisons - Subtask A Load * Subtask B Load - for Subtask A Recall RT [s] in Experiment 4

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-0.126	0.160	27.0	-0.787	1.000
		-	2 [high]	2 [low]	-0.409	0.218	27.0	-1.872	0.288
	3 [high]	-	2 [high]	3 [high]	-0.434	0.179	27.0	-2.430	0.088
2 [high]	2 [low]	-	2 [high]	3 [high]	-0.151	0.204	27.0	-0.739	1.000

5.2.6 Subtask A Presentation RT

Similar to Experiment 1, because of the very nature of the structure of an episode, mA happens before mB, therefore it is only meaningful to not to expect an effect of subtask B WM load on subtask A presentation RT. Again, I will provide the whole picture for the sake of completion. The subtask A presentation reaction time was calculated as the mean amount of time participants took to memorize the lines shown in the screen mA, (the time between the mA started and participant made the button response to move onto the next screen), in seconds. Table 5.16 summarizes the descriptive statistics for subtask B presentation RT. 2x2 Repeated Measures ANOVA (Table 5.17) revealed that only the main effect of subtasks A load was significant (Subtask A Load: $F(1,27) = 34.31, p < 0.001, \eta_p^2 = 0.56$), (Subtask B Load: $F(1,27) = 0.07, p = 0.79, \eta_p^2 = 0.003$), (Interaction: Subtask A Load x Subtask B Load: $F(1,27) = 1.46, p = 0.24, \eta_p^2 = 0.051$).

Pairwise comparisons with Bonferroni correction (Table 5.18) revealed that when subtask B WM load is kept constant between the groups the mean differences between groups were significant ([1,2] vs [2,2]: Mean Difference = -2.24, SE = 0.44, $t(27) = -5.12, p < 0.001$) ([1,3] vs [2,3]: Mean Difference = -1.79, SE = 0.34, $t(27) = -5.32, p < 0.001$). When subtask A WM load is kept constant, on the other hand, the comparisons of [1,2] vs [1,3] (Mean Difference =

-0.19, SE = 0.17, $t(27) = -1.12$, $p \approx 1.00$) and [2,2] vs [2,3] (Mean Difference = 0.26, SE = 0.27, $t(27) = 0.96$, $p \approx 1.00$) were not significant, as logically expected.

Figure 5.2.f illustrates these results.

Table 5.16. Descriptives for Subtask A Presentation RT [s] in Experiment 4

	1_2	1_3	2_2	2_3
N	28	28	28	28
Missing	0	0	0	0
Mean	3.11	3.30	5.35	5.09
Std. error mean	0.270	0.362	0.665	0.569
95% CI mean lower bound	2.58	2.59	4.05	3.98
95% CI mean upper bound	3.64	4.01	6.65	6.21
Median	2.65	2.81	4.37	4.75
Mode	1.27 ^a	1.14 ^a	1.13 ^a	1.29 ^a
Standard deviation	1.43	1.91	3.52	3.01
Minimum	1.27	1.14	1.13	1.29
Maximum	5.94	8.51	14.6	12.7
Skewness	0.732	1.45	0.996	0.838
Std. error skewness	0.441	0.441	0.441	0.441

^a More than one mode exists, only the first is reported

Table 5.17. Within Subjects Effects for Subtask A Presentation RT [s] in Experiment 4

	Sum of Squares	df	Mean Square	F	p	η^2_G	η^2	η^2_p
Subtask A Load	114.0718	1	114.0718	34.3096	< .001	0.135	0.134	0.560
Residual	89.7689	27	3.3248					
Subtask B Load	0.0332	1	0.0332	0.0735	0.788	0.000	0.000	0.003
Residual	12.1828	27	0.4512					
Subtask A Load * Subtask B Load	1.4020	1	1.4020	1.4617	0.237	0.002	0.002	0.051
Residual	25.8978	27	0.9592					

Note. Type 3 Sums of Squares

Table 5.18. Post Hoc Comparisons -Subtask A Load * Subtask B Load - for Subtask A Presentation RT [s] in Experiment 4

Comparison					Mean Difference	SE	df	t	p _{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
1 [low]	2 [low]	-	1 [low]	3 [high]	-0.189	0.169	27.0	-1.122	1.000
		-	2 [high]	2 [low]	-2.242	0.438	27.0	-5.117	< .001
	3 [high]	-	2 [high]	3 [high]	-1.795	0.338	27.0	-5.315	< .001

Table 5.18. Post Hoc Comparisons -Subtask A Load *Subtask B Load - for Subtask A Presentation RT [s] in Experiment 4

Comparison					Mean Difference	SE	df	t	P _{bonferroni}
Subtask A Load	Subtask B Load		Subtask A Load	Subtask B Load					
2 [high]	2 [low]	-	2 [high]	3 [high]	0.258	0.269	27.0	0.960	1.000

5.3 Discussion

Let us first start with the results regarding within-subtask interference. Indication of within-subtask interference would be that a subtask's load would affect its performance. This effect was present in subtask B accuracy, subtask A accuracy, and subtask A presentation RT. Other measures of performance did not show any within-subtask load interference. Across-subtasks load interference, on the other hand, was present for subtask B accuracy, and for any other measures.

The manipulation that we had seemed to lead to disappearing of within-subtask load effects to a greater extent in Experiment 4 in RT measures. This finding is hard to explain with any WM theory including the programs-view we suggested here. One might claim that the design as such was not sensitive enough to make reaction time measurements.

Let's move to the accuracy related findings. Subtask A accuracy showed only within-subtask load effects whereas subtask B accuracy showed both within-subtask and across-subtasks load effect. From task-programs point of view, our manipulation in Experiment 4 arguably led to subtask related programs being more similar to each other compared to Experiment 1. If this assumption holds, one possible explanation of the accuracy related findings might be that this increased similarity between subtask related programs led to an increased across-subtasks interference in Experiment 4, compared to Experiment 1. Despite that the interference is well-established in causing decreased WM performance

and WMC limitations, the established accounts focus on the within-task interference i.e. the interference happening in between the items that belong to the same task set. The across-subtask interference in this current thesis, however, refers to an interference that might occur across different subtasks and their corresponding WM information.

One might also argue that the manipulation that we did in Experiment 4 in terms of equating the presentation mode of the stimuli in both subtasks led to individual items' being held in WM to be more similar, and not necessarily to the subtask-related programs' being more similar. Nonetheless, an objection to this argument might be that increased similarity amongst the items in the WM should have at least led to increased within-subtask interference. Contrary to this prediction, within-subtask interference observed in Experiment 1 disappeared in Experiment 4. Despite all these, it is fair to say that the current design of Experiment 4 can not help us to disambiguate between the two alternatives proposed.

One crucial point in current findings was that across-subtasks interference was observed only during the times both of the subtask related programs are actively maintaining and processing WM information (i.e. during subtask B recall), and not during subtask A recall. In the current hierarchical organization of subtask A and subtask B, during the subtask B recall, both of the subtask programs are still relevant to the completion of higher task episode related goals and task episode program, whereas, when the time for subtask A recall arrives, subtask B related program is no longer goal relevant or active [and arguably removed from the

cognitive space to be instantiated again in the next episode right before the mB phase]. When taken together, these explanations might indicate that across-task program interference is possible when hierarchically organized subtask related programs are concurrently goal-relevant and when the contextual cues regarding their identities bear similarities. This implies that during subtask B recall, the contextual cues that would selectively access to the WM of subtask B program were not distinctive enough to selectively cue the retrieval of the information from the target WM (from the target subtask related program) in Experiment 4 compared to Experiment 1. Such an explanation is pretty similar to ‘interference by confusion’, but this time interference occurs between the concurrently executed subtask related programs, and their corresponding WMs. This might in turn imply that when more than one goal-directed program is actively maintained, they might be competing for attentional selection in response to retrieval cues, and when retrieval cues are not selective enough (i.e. when programs are similar enough) a between-programs interference might happen.

Current findings regarding the accuracy measures do not seem to imply the presence of interference by feature overwriting across subtasks. Feature overwriting between subtask A and subtask B programs or WMs would lead to across-subtask load effects in both subtask A and subtask B. In fact, evidence for feature overwriting, as a mechanism of interference in WM, is found for verbal material, but not for visual material in previous studies, as well [92].

In conclusion, the results of Experiment 4, when compared to the findings from Experiment 1, suggest that interference might occur between concurrently held

and active task related programs' WMs in occasions where their identities are more similar than distinct, at least for accuracy measures.

Chapter 6

General Discussion

6.1 Implications

1. *WM information of hierarchically organized tasks can be maintained separately as part of their separate task-related programs.* In Experiment 1 and Experiment 2, identical WM (sub)tasks of which WM loads are independently manipulated are interleaved. The performance in a given subtask was affected only by the WM load of the corresponding task, and not by the WM load of the concurrently executed subtask. In other words, when multiple subtask-related programs are organized into an extended episode, within-subtask load effects were present whereas across-subtask load effects were not. These results suggested that WM information of a given subtask could be maintained in its subtask-related program separately from the WMs of other concurrently held subtask-related programs.

2. *Maintaining WM information of different task-related programs in their corresponding WMs separately might be a new way of circumventing WMC limits.* If WM information of separate tasks are maintained in their corresponding task-related programs, this could be another way of circumventing WM limits which is generally cited as being four items for visual WM [30]. Maintenance of WM information as parts of separate programs, just like chunking a set of information, might circumvent the item limits beyond these four items.

Remember that the main motivational source of this thesis was the observation that in everyday life we maintain more information in WM than any of WMC views would suggest, and we are able to deal with multiple purposive tasks at once, maintain and process loads of information for them, and manage to control seemingly hierarchical behavior underlying those tasks flexibly. WMC that is portrayed by the findings of this thesis might be a way of showing how this control of WM information that should exceed the limits suggested by the literature is managed: Information that belongs to different goal-directed task-related programs are maintained separately as parts of their programs, and this circumvents WMC limits beyond what is usually suggested by the literature. This is the suggestion of this thesis to the relationship between WMC and hierarchical control of WM by goal-directed task entities.

3. Encoding of information to separate WMs of goal-directed task-related programs might require attentional control of WM, perhaps by the mechanism of consolidation. Encoding of WM information to separate subtask related programs is suggested to require a attentional time-based mechanism in Experiment 3. And, one possibility that arises from the collective findings of Experiment 3 is that simultaneously presented stimuli that belong to a subtask are consolidated together in a time and attention-dependent manner to create a single entity that is closely tied to its corresponding task program, and they might be maintained in it separately from the other subtask's WM information.

4. Increased similarity between subtasks might lead to increased interference between the WMCs of information belonging to separate but concurrent task-related programs. Increasing the similarity between the demands of concurrent tasks led to disappearance of the interference-free execution of concurrent but separate subtasks in Experiment 4, crucially during the times where both subtasks were still goal-relevant to the accomplishment of higher episode-related goals. And, interference-free maintenance of concurrent hierarchically organized task related programs might require some extent of dissimilarity of either task entities that subsumes the execution of the episodes or task information.

5. Attentional control might be mediated by hierarchical control and goal-directed task-episode-related programs. As mentioned in the general introduction, tasks and goals are not just mere linguistic representations of declarative information about the behavior or cognition, rather they are cognitive entities that control the cognitive processes until their completion is arrived. We are referring to these goal-directed cognitive entities as ‘task episode related programs’ [68], or ‘task-related programs’ and any temporally extended sequence of behavior will be subsumed by these programs. In the context of current experiments, too, an extended task episode consisted of subtask A and subtask B with their unique goal states, and hence their own task-related programs will subsume the sequential behavior, and even attentional control processes should be instantiated through the mediation of these goal directed entities in a given task episode. From this perspective, attentional-control-related functions such as the selection of WM information to be held from all the available set of perceptual

and LTM information [e.g. selection of oriented lines' orientation information in relation with their position on the screen and color to be used as retrieval cues] along with consolidation of these information, the retrieval of the relevant information from LTM into WM, and removal of the information that is no longer goal relevant are controlled by these task-related programs. Once they are established and initiated, task-related programs control attentional processes and hierarchical behavior in an orderly and sequential manner.

There are several findings in the current thesis that might suggest that the attentional control of WMC is mediated by the subsuming task-episode related programs.

First finding as such is that in Experiment 2, subtask A recall RT was higher for the high load episodes compared to the low load episode, even though subtask A recall accuracies were the same across all episodes. One suggestion from this finding is that there was an attentional selection in the level of subtask-program causing a competition between the actively maintained subtasks to be retrieved to the attentional focus. Once a subtask has been selectively accessed, however, its content was accessed without any interference from the concurrent load, since the attentional processes are now recruited only by the WM operations within the accessed WM store.

The second finding that might be relevant here is that in Experiment 4, across-subtasks interference has only been observed when both subtasks were still goal-relevant to the accomplishment of higher task-episode-related goals, and

hence active. This finding, also, indicates that the competition for attentional selection between the subtask programs (or between the items belonging to them) was on as long as the subtask-related programs were still a part of the task-episode related programs. When the completion of larger task-episode related program no longer necessitated the completion of one of the subtasks (i.e. after a subtask is completed, and only one is left to be completed in a given episode), there was no such across-subtasks WMC interference.

When taken together, these findings might indicate that attentional control of WM information might be mediated by the larger hierarchical structure of the task-episode in a way that it is performed on different levels of the task-episode related program. When there are more than one actively maintained WMs, first, the relevant subtask-related program and corresponding WM is accessed, and then the target within the contents of the accessed WM is retrieved. In conclusion, attentional control of WM to be selected for further processing might be mediated by the task entities in accordance with the levels [or subtasks] they consist of.

6.2. Further Discussion

To our knowledge, the current thesis is the first study to ever investigate this relationship between WMC and the control instantiated by hierarchical task entities. In addition, the finding of interference-free execution of concurrently executed subtasks is a novel finding in the WMC literature.

The extended task paradigm that is introduced in this thesis is a novel paradigm to investigate WMC-related issues. It differs from the dual-task or complex-span

WM measures that are used widely in the literature in several ways; current task involves two identical subtasks that are interleaved, both of these subtasks are WM tasks, and WMC investigations are done based on continuous stimulus and response space. We believe the current task is a good alternative to the dual-task paradigms when it comes to investigating WMC and how it is controlled by hierarchically organized task entities.

When there is a detrimental effect of the distractor task on the withheld task's performance in dual-task studies, it is concluded that the processing and storage share a common resource [31]. However, there are some previous studies in which such a distractor or a concurrent WM demand did not affect the withheld task performance [32; 33; 34]. Still, note that in these studies the design of the experiment only cares for a unidirectional effect which is the effect of withheld task on the distractor task performance, and uses only one WM task (the second task is not a WM task but, for example, a mental arithmetic task in 34) in their dual task paradigm. Moreover, the WM tasks that are used in these studies require simple maintenance and recall of information without any need to manipulate it in accordance with the task demands. Therefore, the findings of the current thesis makes separate and meaningful contributions to these previous studies.

Sustained neural activity in prefrontal cortex (PFC) during the delay periods in WM tasks has long been taken as a mark of WM and maintenance of WM representations. However, lately, it is suggested that the decodability of WM information from the neuroimaging data might depend on whether the information is high-priority or functionally goal-relevant at a given moment [93]. It is suggested that the representations in WM that are currently and temporarily

goal-irrelevant might take up a latent state indicated by a baseline activity, and such a state would sign that the attention is withdrawn from them. A storage mechanism consisting of such functionally latent states is called ‘the activity silent WM’ [94]. Short-term synaptic plasticity is suggested as a mechanism for such a storage. Studies suggest that WM items that are low-priority or temporarily goal-irrelevant might become behaviorally silent; they might not affect attention-based ongoing processing [95]. For example, in some studies high-priority items in WM affected the visual search whereas the low-priority or currently goal-irrelevant items (that are perhaps activity silent) did not [96; 97; 95]. In current thesis, too, WM information that is encoded in relation to a subtask related program might be taking such a latent state, and during which the times it stays goal-irrelevant, it might not affect WMC by a mechanism similar to activity-silent WM. How, on the other hand, activity silent WM information is coded to achieve that is unclear, and whether the current paradigm leads to such encoding should be further investigated.

6.3 Limitations

Current study did not really focus on whether the performance differs across subtask A and subtask B under the same load conditions. For example, in Experiment 1, if subtask A information is actively kept in background in relation to its own subtask related program during the execution of subtask B and subtask A and subtask B loads did not interfere with each other, one question follows would be whether subtask A performance differs from subtask B performance of the same load. Answering this question systematically would provide additional beneficial information about the nature of the WM maintenance by the

task-episode related programs, specifically whether it is subject to interference from other cognitive processes or decay. However, for the current thesis this question was not central, and the experimental paradigms were not equipped to investigate such a matter. Future studies might investigate the effect of time passage on the subtask A performance and its interaction with other capacity-limitation issues in a design similar to the current one.

RT related results are somehow more inexplicable (compared to the accuracy related results) within the framework that is suggested in this thesis regarding the relationship between the hierarchical control of WM and WMC. This might reflect a limitation of the current task paradigms in terms of sensitivity in measuring the RTs especially since the recalls were self-paced (except for the Experiment 4), and changing the orientation of the probe in recall phases required participants to sometimes go back and forth to report what they had in their mind, eventually spend more time in recall screen than what their recall took. In addition, the initial orientation of the probe is randomly determined, a pseudorandomization in terms of how long does it take to rotate the probe from its original position to the target position could have been made. Future studies should replicate the results in settings and paradigms that are more sensitive to the RT measurement.

In Experiment 4, instead of systematically testing the effect of varied stimulus onset asynchronies (SOA) on the performance, we used a cut-off SOA of 700 ms. Although this was an informed decision based on the previous literature, using varied SOAs in a systematic way instead of using a single fixed SOA could have

led to a more holistic understanding of the time constraints on the encoding of WM information into separate task related program stores. Future studies might address this issue.

Decay of information as a hypothesis for why WM is capacity limited could not be addressed with the current designs. We cannot argue that WM information of the subtask information that is maintained in relation to its corresponding subtask related program was not subject to decay because we did not have a test condition where we manipulated the time between subtask A memorization and recall, and then made the necessary analysis. Future studies might test to what extent the temporarily goal-irrelevant information that is kept in its own task-related program.

6.4 Future Directions

It is obvious that some WM tasks that have a hierarchical structure to them might not show signs of separate WMs for separate subtask related programs. Although the results were promising in Experiment 1 and Experiment 2 in terms of creating instances where different task programs can have their own WMs, the exact aspects of the current extended hierarchical WM task that would be responsible for driving such results are unclear. One might claim task order information takes up some WMC, in addition to all the WM information that is to be used in concurrent processing. Task order is a higher-level information that would be derived from the hierarchical execution of the given task, just like the hierarchical structure of our task here. In the current extended task, the subtask demands were almost exactly the same, and the rules governing both of the subtasks were

identical, except for the rules governing the overall hierarchical execution of them. Although at the first glance, this seems like a disadvantage (i.e. it may increase task interference due to attentional processes, or since both tasks are relying on the same ‘slave system’ [16]), in fact, it is possible that by having such a simple rule (i.e. memorize the orientation of lines and retrieve them when prompted) to be followed throughout both of the subtasks might have lowered the cognitive load due to the task structure itself. It might even have eased the process of construing the hierarchical task entity as a single task program. Another possibility is that, in addition to being similar to each other to ease the construal, the task related programs of subtasks were perfectly distinguishable with small amounts of processing. Since the subtask demands were so similar, the contextual information that helps to distinguish one subtask from another might have had the chance to be elevated even with the smallest allocation of capacity-limited attentional or non-attentional WM resources, and their similarities could always stay in the background where resources are non-capacity limited, and never needed to be brought into attentional focus during the ongoing processing. Therefore, the subtask related programs (or contexts) were consistently and strongly separate from each other, which in turn resulted in their own corresponding WMs. The effect of the exact hierarchical nature of a given task episode should be investigated further in the future.

‘Tasks’ and ‘goals’ that are hierarchically stored in our memory might actually be remaining memories of the tasks and goals that we engaged with in the past. Farooqui and Manly [68] also pointed out the possibility that the stronger previous task execution memories that are used extensively in the past might be

utilized to control the execution of the current task program. Such an understanding is also important when it comes to the relationship between practice prior to the task and the development of the hierarchical structure or task program. There are some accounts that suggest that prior to any learning or practice, the hierarchical nature of the sequence of actions should be readily available to the doer [98]. At the same time it is obvious that the learning and practice should also provide the doer with the knowledge of sequence [70]. Such an investigation could be done in the future regarding the effect of practice or familiarity of the hierarchical task structure in the development of such hierarchical WM control.

Investigations on how the behavioral results obtained in this thesis in terms of the interference-free maintenance of WM information would be reflected in neuroimaging and psychophysiological data might be another future direction. As stated, especially the similarities between activity-silent WM accounts and the current findings remains to be investigated.

BIBLIOGRAPHY

- [1] Postle, B. R., & Oberauer, K. (2022). Chapter to appear in M.J. Kahana and A.D. Wagner (Eds.), *The Oxford Handbook of Human Memory*. DOI: <https://doi.org/10.31234/osf.io/963kf>
- [2] Velichkovsky, B. B. (2017). Consciousness and working memory: Current trends and research perspectives. *Consciousness and Cognition*, 55, 35-45. DOI: <https://doi.org/10.1016/j.concog.2017.07.005>
- [3] 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. DOI: https://doi.org/10.1044/2018_LSHSS-17-0114
- [4] Miller, G. A., Galanter, E., Pribram, K. H. (1960). *Plans and the Structure of Behavior*. New York: Henry Holt and Company.
- [5] 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. DOI: <https://psycnet.apa.org/doi/10.1037/0096-3445.128.3.309>
- [6] Gold, J. M., Barch, D. M., Feuerstahler, L. M., Carter, C. S., MacDonald, A.

- W., Ragland, J. D., Silverstein, S. M., Strauss, M. E., & Luck, S. J. (2018). Working memory impairment across psychotic disorders. *Schizophrenia Bulletin*, 45(4), 804–812. DOI: <https://doi.org/10.1093/schbul/sby134>
- [7] Ortega, R., López, V., Carrasco, X., Escobar, M. J., García, A. M., Parra, M. A., & Aboitiz, F. (2020). Neurocognitive mechanisms underlying working memory encoding and retrieval in Attention-Deficit/Hyperactivity Disorder. *Scientific reports*, 10(1), 1-13. DOI: <https://doi.org/10.1038/s41598-020-64678-x>
- [8] Stopford, C. L., Thompson, J. C., Neary, D., Richardson, A. M., & Snowden, J. S. (2012). Working memory, attention, and executive function in Alzheimer's disease and frontotemporal dementia. *Cortex*, 48(4), 429-446. DOI: <https://doi.org/10.1016/j.cortex.2010.12.002>
- [9] Ramos, A. A., & Machado, L. (2021). A comprehensive meta-analysis on short-term and working memory dysfunction in Parkinson's disease. *Neuropsychology Review*, 31(2), 288-311.
- [10] Shan, Y., Jia, Y., Zhong, S., Li, X., Zhao, H., Chen, J., ... & Wang, Y. (2018). Correlations between working memory impairment and neurometabolites of prefrontal cortex and lenticular nucleus in patients with major depressive disorder. *Journal of affective disorders*, 227, 236-242. DOI: <https://doi.org/10.1016/j.jad.2017.10.030>
- [11] Cowan, N. (2017). The many faces of working memory and short-term

- storage. *Psychonomic bulletin & review*, 24(4), 1158-1170.
- [12] 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. DOI: <https://psycnet.apa.org/doi/10.1037/0278-7393.28.3.411>
- [13] Lewis-Peacock, J. A., Kessler, Y., & Oberauer, K. (2018). The removal of information from working memory. *Annals of the New York Academy of Sciences*, 1424(1), 33-44. DOI: <https://doi.org/10.1111/nyas.13714>
- [14] Öztekin, I., & Cowan, N. (2015). Representational states in memory: where do we stand?. *Frontiers in Human Neuroscience*, 9, 453. DOI: <https://doi.org/10.3389/fnhum.2015.00453>
- [15] LaRocque, J. J., Lewis-Peacock, J. A., and Postle, B. R. (2014). Multiple neural states of representation in short-term memory? It's a matter of attention. *Front. Hum. Neurosci.* 8:5. DOI: 10.3389/fnhum.2014. 00005
- [16] Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.
- [17] Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory*. Vol. 2. New York: Academic Press, 1968. Pp. 89-195.

- [18] Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, 225(2), 82-91.
- [19] Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, 4(11), 417-423. DOI: [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- [20] Baddeley, A. (2002). Fractionating the central executive. *Principles of frontal lobe function*, 246-260.
- [21] Cowan, N. (2008). What are the differences between long-term, short-term, and working memory?. *Progress in brain research*, 169, 323-338. DOI: [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- [22] Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. Published online by Cambridge University Press.
- [23] Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological bulletin*, 104(2), 163.
- [24] Hebb D. O. (1949). *The Organisation of Behaviour*. New York, NY: John Wiley & Sons.

- [25] Langille, J. J., & Brown, R. E. (2018). The synaptic theory of memory: a historical survey and reconciliation of recent opposition. *Frontiers in systems neuroscience*, 12, 52. DOI: <https://doi.org/10.3389/fnsys.2018.00052>
- [26] Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, 75(6), 522-536.
- [27] Cowan, N. (1999). An Embedded-Processes Model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press.
- [28] James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology* (Vol. 1, No. 2). London: Macmillan.
- [29] Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences*, 24(1), 87-114.
- [30] Cowan, N. (2012). *Working memory capacity*. Psychology press.
- [31] Mathy, F., Chekaf, M., & Cowan, N. (2018). Simple and complex working memory tasks allow similar benefits of information compression. *Journal of Cognition*, 1(1). DOI: <https://doi.org/10.5334%2Fjoc.31>
- [32] Foos, P. W., & Wright, L. (1992). Adult age differences in the storage of

- information in working memory. *Experimental Aging Research*, 18(2), 51-57.
- [33] Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General*, 112(2), 240.
- [34] Oberauer, K., Demmrich, A., Mayr, U., & Kliegl, R. (2001). Dissociating retention and access in working memory: An age-comparative study of mental arithmetic. *Memory & Cognition*, 29(1), 18-33. DOI: <https://doi.org/10.3758/BF03195737>
- [35] Garavan, H. (1998). Serial attention within working memory. *Memory & cognition*, 26(2), 263-276.
- [36] Oberauer, K. (2013). The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in human neuroscience*, 7, 673. DOI: <https://doi.org/10.3389/fnhum.2013.00673>
- [37] Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity?. *Psychological bulletin*, 142(7), 758. DOI: <https://psycnet.apa.org/doi/10.1037/bul0000046>
- [38] Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279-281.

- [39] Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, 13(1-2), 187-215. DOI: <https://doi.org/10.1080/09541440042000278>
- [40] Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2), 81.
- [41] Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current opinion in neurobiology*, 20(2), 177-182. DOI: <https://doi.org/10.1016/j.conb.2010.03.005>
- [42] Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, 17(3), 347-356.
- [43] Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of vision*, 4(12), 11-11.
- [44] Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of vision*, 9(10), 7-7.
- [45] Gorgoraptis, N., Catalao, R. F., Bays, P. M., & Husain, M. (2011). Dynamic

- updating of working memory resources for visual objects. *Journal of Neuroscience*, 31(23), 8502-8511. DOI: <https://doi.org/10.1523/JNEUROSCI.0208-11.2011>
- [46] Van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological review*, 121(1), 124. DOI: <https://psycnet.apa.org/doi/10.1037/a0035234>
- [47] Ku, Y. (2018). Selective attention on representations in working memory: cognitive and neural mechanisms. *PeerJ*, 6, e4585. DOI: <https://doi.org/10.7717/peerj.4585>
- [48] Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General*, 130(2), 169. DOI: <https://psycnet.apa.org/doi/10.1037/0096-3445.130.2.169>
- [49] Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067), 500-503. DOI: <https://doi.org/10.1038/nature04171>
- [50] 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. DOI: <https://doi.org/10.1111/nyas.13943>

- [51] 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. DOI: <https://doi.org/10.1016/j.jml.2014.12.004>
- [52] Ricker, T. J. (2015). The role of short-term consolidation in memory persistence. *AIMS Neuroscience*, 2(4), 259-279. DOI: <https://doi.org/10.3934/Neuroscience.2015.4.259>
- [53] Ricker, T. J., Nieuwenstein, M. R., Bayliss, D. M., & Barrouillet, P. (2018). Working memory consolidation: insights from studies on attention and working memory. *Annals of the New York Academy of Sciences*, 1424(1), 8-18. DOI: <https://doi.org/10.1111/nyas.13633>
- [54] 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. DOI: <https://psycnet.apa.org/doi/10.1037/a0034301>
- [55] Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*, 21(6), 413-419. DOI: <https://doi.org/10.1177/0963721412459513>
- [56] Vergauwe, E., & Cowan, N. (2014). A common short-term memory retrieval

rate may describe many cognitive procedures. *Frontiers in Human Neuroscience*, 8, 126. DOI: <https://doi.org/10.3389/fnhum.2014.00126>

[57] Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory?. *Annals of the new York Academy of Sciences*, 1424(1), 19-32. DOI: <https://doi.org/10.1111/nyas.13616>

[58] 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. DOI: <https://doi.org/10.1016/bs.plm.2017.03.005>

[59] Johnson, M. R., & Johnson, M. K. (2009). Toward characterizing the neural correlates of component processes of cognition. *Neuroimaging of human memory: Linking cognitive processes to neural systems*, 169-194.

[60] Ecker, U. K., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language*, 74, 77-90. DOI: <https://doi.org/10.1016/j.jml.2013.09.003>

[61] Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: attending to one item frees capacity. *Journal of*

Experimental Psychology: Human Perception and Performance, 40(3), 1237.

DOI: <https://psycnet.apa.org/doi/10.1037/a0036331>

[62] Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the Time-Based Resource-Sharing theory. *Psychonomic bulletin & review*, 18(1), 10-45. DOI: <https://doi.org/10.3758/s13423-010-0020-6>

[63] Oberauer, K. (2018). Removal of irrelevant information from working memory: Sometimes fast, sometimes slow, and sometimes not at all. *Annals of the New York Academy of Sciences*, 1424(1), 239-255. DOI: <https://doi.org/10.1111/nyas.13603>

[64] Ricker, T. J., Spiegel, L. R., & Cowan, N. (2014). Time-based loss in visual short-term memory is from trace decay, not temporal distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1510. DOI: <https://psycnet.apa.org/doi/10.1037/xlm0000018>

[65] Oberauer, K. (2001). Removing irrelevant information from working memory: a cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 27(4), 948. DOI: <https://psycnet.apa.org/doi/10.1037/0278-7393.27.4.948>

[66] Barto, A. G., & Mahadevan, S. (2003). Recent advances in hierarchical

reinforcement learning. *Discrete event dynamic systems*, 13(1), 41-77. DOI:
<https://doi.org/10.1023/A:1022140919877>

[67] Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, 113(3), 262-280. DOI:
<https://doi.org/10.1016/j.cognition.2008.08.011>

[68] Farooqui, A. A., & Manly, T. (2019). We do as we construe: extended behavior construed as one task is executed as one cognitive entity. *Psychological Research*, 83(1), 84-103.

[69] Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: switching tasks in sequences. *Journal of Experimental Psychology: General*, 135(4), 623. DOI: <https://psycnet.apa.org/doi/10.1037/0096-3445.135.4.623>

[70] Botvinick, M., & Plaut, D. C. (2004). Doing Without Schema Hierarchies: A Recurrent Connectionist Approach to Normal and Impaired Routine Sequential Action. *Psychological Review*, 111(2), 395–429. DOI:
<https://doi.org/10.1037/0033-295X.111.2.395>

[71] Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, 104(1), 3–65.

- [72] Kruglanski, A. W., & Kopetz, C. (2009). What is so special (and non-special) about goals? A view from the cognitive perspective. In H. G. Gordon & B. Moskowitz (Eds.), *The psychology of goals* (pp. 27–55). New York: Guilford Press.
- [73] Farooqui, A. A., Mitchell, D., Thompson, R., & Duncan, J. (2012). Hierarchical organization of cognition reflected in distributed frontoparietal activity. *Journal of Neuroscience*, 32(48), 17373-17381.
- [74] Farooqui, A. A., & Manly, T. (2018). When attended and conscious perception deactivates fronto-parietal regions. *Cortex*, 107, 166-179.
- [75] The jamovi project (2021). *jamovi* (Version 2.2) [Computer Software]. Retrieved from <https://www.jamovi.org>
- [76] Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, 51(1), 195-203. DOI: <https://doi.org/10.3758/s13428-018-01193-y>
- [77] Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in quantitative methods for psychology*, 1(1), 42-45.

- [78] Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2), 61-64.
- [79] R Core Team (2021). R: A language and environment for statistical computing. (Version 4.0) [Computer Software]. Retrieved from <https://www.jamovi.org>. Retrieved from <https://cran.r-project.org>. (R packages retrieved from MRAN snapshot 2021-04-01).
- [80] Singmann, H. (2018). afex: Analysis of Factorial Experiments. [R package]. Retrieved from <https://cran.r-project.org/package=afex>.
- [81] Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current directions in psychological science*, 21(3), 164-169. DOI: <https://doi.org/10.1177/0963721412444727>
- [82] 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. DOI: <https://psycnet.apa.org/doi/10.1037/0096-1523.32.6.1436>
- [83] Jarjat, G., Hoareau, V., Plancher, G., Hot, P., Lemaire, B., & Portrat, S. (2018). What makes working memory traces stable over time?. *Annals of the New York Academy of Sciences*, 1424(1), 149-160. DOI: <https://doi.org/10.1111/nyas.13668>

- [84] Knöchel, C., Oertel-Knöchel, V., Bittner, R., Stäblein, M., Heselhaus, V., Prvulovic, D., ... & Linden, D. E. (2015). Consolidation time affects performance and neural activity during visual working memory. *Psychiatry Research: Neuroimaging*, 231(1), 33-41. DOI: <https://doi.org/10.1016/j.psychresns.2014.10.025>
- [85] Dell'Acqua, R., Sessa, P., Toffanin, P., Luria, R., & Jolicoeur, P. (2010). Orienting attention to objects in visual short-term memory. *Neuropsychologia*, 48(2), 419-428. DOI: <https://doi.org/10.1016/j.neuropsychologia.2009.09.033>
- [86] Wyble, B., Bowman, H., & Nieuwenstein, M. (2009). The attentional blink provides episodic distinctiveness: sparing at a cost. *Journal of experimental psychology: Human perception and performance*, 35(3), 787. DOI: <https://psycnet.apa.org/doi/10.1037/a0013902>
- [87] Wyble, B., Potter, M. C., Bowman, H., & Nieuwenstein, M. (2011). Attentional episodes in visual perception. *Journal of Experimental Psychology: General*, 140(3), 488. DOI: <https://psycnet.apa.org/doi/10.1037/a0023612>
- [88] Portrat, S., Guida, A., Phénix, T., & Lemaire, B. (2016). Promoting the experimental dialogue between working memory and chunking: Behavioral data and simulation. *Memory & cognition*, 44(3), 420-434. DOI: <https://doi.org/10.3758/s13421-015-0572-9>

- [89] Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2013). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, Advance online publication. DOI: 10.1037/a0034221
- [90] Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105, 761-781. DOI: <https://psycnet.apa.org/doi/10.1037/0033-295X.105.4.761-781>
- [91] 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, 627-652. DOI: <https://doi.org/10.1016/j.jml.2006.08.005>
- [92] Jünger, E., Kliegl, R., & Oberauer, K. (2013). No evidence for feature overwriting in visual working memory. *Memory*, 22, 374-389. DOI: 10.1080/09658211.2013.794243
- [93] Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, 319(5869), 1543-1546. DOI: <https://doi.org/10.1126/science.1150769>
- [94] Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of cognitive neuroscience*, 24(1), 61-79. DOI: https://doi.org/10.1162/jocn_a_00140

- [95] Mallett, R., & Lewis-Peacock, J. A. (2018). Behavioral decoding of working memory items inside and outside the focus of attention. *Annals of the New York Academy of Sciences*, 1424(1), 256-267. DOI: <https://doi.org/10.1111/nyas.13647>
- [96] van Moorselaar, D., Theeuwes, J., & Olivers, C. N. (2014). In competition for the attentional template: Can multiple items within visual working memory guide attention?. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1450. DOI: <https://psycnet.apa.org/doi/10.1037/a0036229>
- [97] Greene, C. M., Kennedy, K., & Soto, D. (2015). Dynamic states in working memory modulate guidance of visual attention: Evidence from an n-back paradigm. *Visual Cognition*, 23(5), 546-560. DOI: <https://doi.org/10.1080/13506285.2015.1052033>
- [98] Houghton, G., (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish & M. Zock (Eds.), *Current research in natural language generation*. London: Academic Press.