RELEASE FROM PROACTIVE INTERFERENCE AND ITS RELATIONS TO EXECUTIVE FUNCTIONS: A DEVELOPMENTAL STUDY ON TURKISH CHILDREN

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

GÜLTEN ÜNAL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

SEPTEMBER 2008

Approval of the Graduate School of Informatics

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Deniz Zeyrek
Head of Department

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

Assist. Prof. Annette Hohenberger
Supervisor

Examining Committee Members

Assist. Prof. Bilge Say	(METU, COGS)	
Assist. Prof. Annette Hohenberger	(METU, COGS)	
Dr. Mine Mısırlısoy	(METU, PSY)	
Assoc. Prof. Emre Özgen	(Bilkent, PSY)	
Dr. Ceyhan Temürcü	(METU, COGS)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this wok.

Name, Last name : Gülten Ünal

Signature : ________________

ABSTRACT

RELEASE FROM PROACTIVE INTERFERENCE AND ITS RELATIONS TO EXECUTIVE FUNCTIONS: A DEVELOPMENTAL STUDY ON TURKISH CHILDREN

Ünal, Gülten
M. Sc., Department of Cognitive Science
Supervisor: Assist. Prof. Annette Hohenberger

September 2008, 146 pages

The aim of this study was to investigate the development of release from proactive interference (RPI) and its relations with executive working memory functions. 101 primary school children (aged 6-12 years) and 20 young adults (aged 22-30 years) participated in the study. The main task, the Categorical Free Recall Test, comprised 12 items from 3 different categories (animals, fruits, clothes). The purpose of the main task was to examine both the development of the RPI pattern and the categorization ability during childhood. As our results showed, the categorization ability and the RPI pattern were already present in the 1st graders. Although overall memory span increased with age, there was no significant development for the categorization and the RPI effect. For the additional tasks, the Word Span Test

(WST, to measure the phonological WM capacity), the Wisconsin Card Sorting Test (WCST, to measure both the categorization ability and executive WM functions), and the Listening Span Test (LST, to examine executive and complex WM functions), the results indicated that children also improved with age. Overall memory capacity in the main task was best predicted by the WST; however, memory of serial position was best predicted by the LST. These findings are in accordance with the view that the WST measures the phonological working memory span, whereas the LST measures complex working memory and executive functions. The comparisons between the adult and the child sample revealed that except for the RPI pattern adults were better on all tasks than the children. The lack of a consistent RPI pattern for the adults may be due to the relatively short stimulus list.

Keywords: Release form Proactive Interference (RPI), Categorical Free Recall Test (CFR-Test), Categorization, Cognitive Development, Executive Functions

ÖZ

İLERİYE DOĞRU BOZUCU ETKİDEN KURTULMA VE BUNUN YÖNETİCİ FONKSİYONLARLA OLAN İLİŞKİLERİ: TÜRK ÇOCUKLARINDA GELİŞİMSEL BİR ÇALIŞMA

Ünal, Gülten
Yüksek Lisans, Bilişsel Bilimler
Tez Yoneticisi: Yrd. Doç. Dr. Annette Hohenberger

Eylül 2008, 146 sayfa

Bu çalışmanın amacı ileriye doğru bozucu etkiden kurtulmanın gelişimini ve bunun diğer yönetici işler bellek fonksiyonlarla olan ilişkilerini incelemektir. Çalışmaya 101 ilkokul öğrencisi (6-12 yaş aralığında) ve 20 genç yetişkin (22-30 yaş aralığında) katılmıştır. Ana test olan Kategorisel Serbest Hatırlama Testi 3 farklı kategoriden (hayvanlar, meyveler, giysiler) 12 adet kelime içermektedir. Bu testin amacı çocukluk dönemi boyunca hem ileriye doğru bozucu etkiden kurtulma örüntüsünün hem de kategorize edebilme yeteneğinin gelişimini incelemektir. Sonuçlarımız kategorize edebilme yeteneğinin ve ileriye doğru bozucu etkiden kurtulma örüntüsünün 1. sınıflarda bile bulunduğunu göstermiştir. Genel olarak hafıza aralığı yaşla beraber artarken, ileriye doğru bozucu etkiden kurtulma faktöründe ve kategorizasyonda önemli bir gelişme olmamıştır. Ek olarak yapılan testlerin (fonolojik işler bellek kapasitesini ölçmek için Kelime Aralığı Testi, hem kategorize edebilme yeteneğini hem de yönetici işler bellek fonksiyonlarını ölçmek için Wisconsin Kart Eşleştirme Testi, yönetici ve karmaşık işler bellek fonksiyonlarını incelemek için Dinleme Aralığı Testi) sonuçları çocukların yaşa bağlı olarak geliştiğini göstermektedir. Ana testteki genel hafıza kapasitesi en iyi KelimeAralığı Testi tarafından tahmin edilmiştir fakat listedeki sıraya uygun olarak hatırlama kapasitesi en iyi Dinleme Aralığı Testi tarafından tahmin edilmiştir. Bu bulgular Kelime Aralığı Testi'nin fonolojik işler bellek aralığını, Dinleme Aralığı Testi'nin de karmaşık işler bellek ve yönetici fonksiyonları ölçtüğü görüşü ile uyumludur. Yetişkin ve çocuk grupları arasındaki karşılaştırmalar, ileriye doğru bozucu etkiden kurtulma örüntüsü hariç yetişikinlerin çocuklara göre her deneyde daha iyi olduklarını ortaya çıkarmıştır. Yetişkinler için tutarlı bir ileriye doğru bozucu etkiden kurtulma örüntüsünün eksikliği nispeten kısa olan uyarıcı listesi sebebiyle olabilir.

Anahtar Kelimeler: İleriye Doğru Bozucu Etkiden Kurtulma, Kategorisel Serbest Hatırlama Testi, Kategorizasyon, Bilişsel Gelişim, Yönetici Fonksiyonlar

To My Little Angle Sinoş

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor Assist. Prof. Dr. Annette Hohenberger for her invaluable support, suggestions, guidance and patience throughout this study. Without her support, this thesis could not have been completed.

Thanks to the Müge Bayramoğlu who allowed me to use the Word Span Test in my study. Also, thanks to Theodore Marinis and Duygu Özge from Reading University for their collaboration on the Listening Span Test. It was a pleasure for me to work with them on this exciting topic.

Special thanks to the Havva Hoca, the director of Yozgat Akdağmadeni Kırlar Köyü İlköğretim Okulu, for giving me the permission to carry out my study at her school and her friendly attitude. I would like to thank the teachers and the director, Tahir Hoca, of Yünalan Köyü İlköğretim Okulu for welcoming me during the experiments. Besides, thanks to my sister, Hatice, for allowing me to carry out the pilot studies in her school, Karaalikaçağı İlköğretim Okulu.

Millions of thanks to my dear parents for their marvelous friendship and patience during this study. I dedicate this thesis to my little sister Sinem who inspired me to study with children. Without her love and empathy, I could never manage to finish this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
CFR-Test	Categorical Free Recall Test
LST	Listening Span Test
PI	Proactive Interference
RPI	Release from Proactive Interference
RST	Reading Span Test
SD	Standard Deviation
WCST	Wisconsin Card Sorting Test
WM	Working Memory
WST	Word Span Test

CHAPTER 0

INTRODUCTION

For the last 30 years working memory has been a highly active research area. Some of the most important studies have been carried out by Baddeley and his colleagues (as surveyed in Baddeley, 2003). The early working memory model introduced by Baddeley and Hitch (1974) maintains that there are two slave systems, the phonological loop and the visuospatial sketchpad, that support the working memory system. These slave systems are controlled by the central executive, also named attentionally-limited control system. According to this three-component model of working memory, the phonological loop was defined as temporary verbal-acoustic storage. The visuospatial sketchpad has the same responsibilities as the phonological loop but for visual storage and manipulation. Finally the central executive was claimed to control our behaviors while being a limited capacity attentional system (For a more detailed account of Baddeley's multi-component working memory model, see the Literature Review in Chapter 1).

The development of working memory has also been studied extensively. Gathercole (1999) found that working memory performance increases rapidly until 8 years of age and then displays a somehow slower development up to 11-12 years of age. Unlike the development of the phonological and the visuospatial working memory, complex working memory in the service of the central executive displays further improvement until 16 years of age. This delayed development may indicate the relatively late development of the related brain areas (i.e., frontal lobes) regarding the complex working memory functions.

The categorization ability is a general ability of humans. Already infants start to categorize objects in the second year of their lives (Younger and Fearing, 1999). However, surprisingly, some animals also display some knowledge of categories. For example, Inoue et al. (2008) found that "rhesus monkeys are able to perform visual discrimination of highly abstract biologically significant categories with better performance in a food-related category than a gender-related one, using two-dimensional visual information" (p. 70). Besides, Dukas and Waser (1994) tested the categorization ability for colors of flowers in bumblebees. They found that bumblebees categorized flowers as rewarding and non-rewarding. Lastly, some animals also do have phonemic categories such as Japanese quail. Kluender et al. (1987) observed that "Japanese quail learned a category for syllable-initial [d] followed by a dozen different vowels" (p. 1195). That is, the categorization ability can be observed both in humans and animals, albeit in different forms.

Related with category knowledge are two memory phenomena: Proactive Interference (PI) and Release from Proactive Interference (RPI). Proactive interference (PI) can be described as follows: Previously learned items proactively interfere with newly learned items. If, however, these stimuli are organized into two sets from different categories, release from proactive interference occurs, that is, the items from the new category are exempt from proactive interference. Release from proactive interference (RPI) can be measured by constructing a categorical short-term memory experiment. In this task, subjects listen to stimuli from different categories, unlike the normal short-term memory experiments in which the items are from the same category (or not organized into categories at all). The subjects are told to remember what they just heard. In a normal free recall experiment, the number of remembered items decreases from the beginning of the items to the end. However, the result of the Categorical Free Recall experiment shows that there is release of proactive interference between the last item of each category and the first item of the next category. Since the items from the first set are from the same category, PI builds up. However, the remaining items are from another category, therefore RPI occurs since these two different item lists do not interfere with each other. Within the second category, PI builds up again, and if this second set is again followed by a new categorical set, then again one could observe the second RPI, and so on.

There are few studies addressing the development of PI and RPI, especially studies with younger children are missing. RPI studies (for adults as well as school children) in the literature are conducted in the following way (e.g. the Peterson & Peterson task, 1959, as cited in Douglas & Corsale, 1977): Firstly, the participants listen to 3 sets of items from the same semantic category (e.g., animals) in which there exist 4 items for each set. Then, in the 4^{th} set they listen to items from a different category (e.g., fruits). That is, there is only one category shift in these experiments. Thus, the build-up of PI is expected till the end of the 3^{rd} set. RPI occurs between the 3^{rd} and the 4^{th} sets since after having listened to the items from the same category, the new category will raise the attention of the participant and its items will be remembered better.

Given the lack of systematic studies on the development of RPI in children, we set out to study the development of categorization and working memory and bring them together in the study of PI and RPI.

In our main experiment, the RPI task, we used two category shifts because we also wanted to know whether PI and RPI would manifest themselves in a regular way across a somewhat more varied sample of items. There were three categories in the task, namely fruits, animals, and clothes. Before the experiment, we collected spontaneous data from subjects in the primary school age on various categories. After some pilot studies on the suitability of those categories, we decided to use the present ones since children at various ages showed relatively good knowledge about these categories. Other possible categories that we had explored before such as colors, vehicles, and furniture had been dismissed for various reasons. The first one, color, has the potential to produce several false memories. The results on the other two categories, vehicles and furniture, had shown that the children only infrequently recalled items from the vehicle category, and even less from the furniture category.

Apart from the main task, the RPI task, we devised a couple of additional measurements: a Word Span Test (WST), the Wisconsin Card Sorting Test (WCST), and a Listening Span Test (LST, for children) as well as a Reading Span Test (RST, for adults). The WST was carried out to measure the phonological memory span of the subjects. The WCST, the LST and the RST are complex working memory tests which measure executive functions (see Chapter 3). All these additional measures were added to investigate the relationship between the memory task and the executive functions of the working memory system.

The WST was constructed by the author, and improved a previous Turkish version used in Bayramoglu and Hohenberger (2005). It includes one-syllabic Turkish words of medium to high frequency. Also, the Turkish Reading Span Test was created by the author, and the test was similar to that of Saito and Miyake (2004). In this experiment, the sentences were selected from school books for children of medium age consisting of widely known facts. The structure of the test is the same as the LST (see Chapter 3).

The Turkish Listening Span Test was the result of the collaboration with Dr. Theodore Marinis (University of Reading), Duygu Özge (University of Reading), Dr. Annette Hohenberger, and Gülten Ünal. The original LST (Pickering and Gathercole, 2001) was translated and adapted for Turkish children.

Our sample was a big rural sample from Yozgat including 101 students. We also conducted control experiments with young adults (METU students) in order to test the differences between the child and the adult group. Thus, the developmental line ranged from 6 to 24 years, approximately. The study aimed to be a detailed study on the development of RPI in Turkey.

The main results of this study about the children were presented in form of posters and talks in Marmaris (5th International Cognitive Neuroscience Meeting, Marmaris, Turkey, 17 - 21 May 2008), Ilgaz (2. Psikoloji Lisansüstü Öğrencileri Kongresi, Ilgaz, Ankara, Türkiye, 26 - 29 Haziran 2008), and Bodrum (10th International Conference on Cognitive Neuroscience, Bodrum, Turkey, September 1st-5th, 2008).

CHAPTER 1

LITERATURE REVIEW

1.1 Working Memory and Working Memory Development

The first working memory model was introduced by Baddeley and Hitch (1974, as cited in Repovs & Baddeley, 2006). It was a multicomponential model in which there are three functional components: the central executive, the phonological loop, and the visuospatial sketchpad. The central executive reflects a limited attentional capacity and is responsible for controlling the other two slave components. In a nutshell, the phonological loop is responsible for storing the phonological information in working memory while the visuospatial sketchpad is assumed to be dedicated to visual and spatial information. Figure 1 shows the original tripartite structure of the first working memory model.

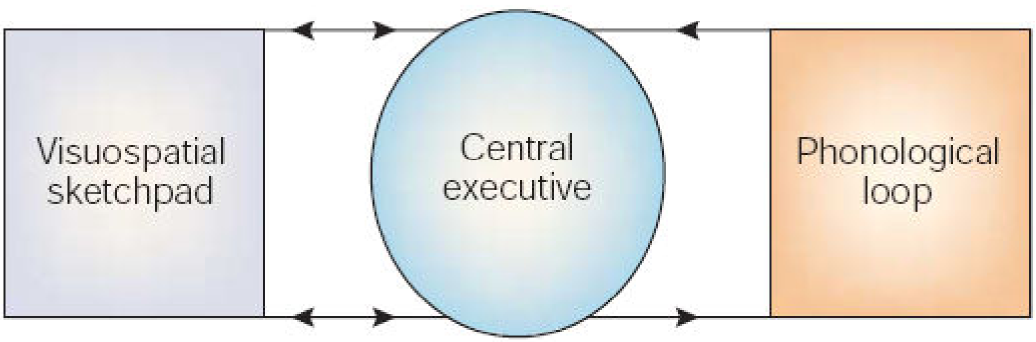


Figure 1 The tripartite working memory model of Baddeley and Hitch (1974). (Baddeley 2003: 830)

Gathercole et al. (2004) defines working memory as follows: "a mental workplace in which information can be stored and processed for brief periods of time in the course of demanding cognitive activities" (p.2). Alternatively, working memory was defined as a system that includes "(a) a store in the form of long-term memory traces active above threshold, (b) processes for achieving and maintaining that activation, and (c) controlled attention" (Engle, Kane, and Tuholski, 1999, p. 104, as cited in Kail & Hall, 2001, p.1). In the first definition, working memory is the functional aspect of short-term memory which is conceived of as a separate memory module. In the second definition, it is a temporarily activated sub-component of long-term memory by means of an active attentional mechanism.

After several studies in the area of working memory, a new component, the episodic buffer, was added to the working memory model (Baddeley, 2000, as cited in Repovs & Baddeley, 2006). It is also assumed to be "a limited capacity store that is capable of multi-dimensional coding, and that allows the binding of information to create integrated episodes" (p. 7). The episodic buffer is associated with the central executive on the one hand and with episodic long term memory on the other hand. Figure 2 shows the multi-componential structure of the recent working memory model of Baddeley and his collaborators.

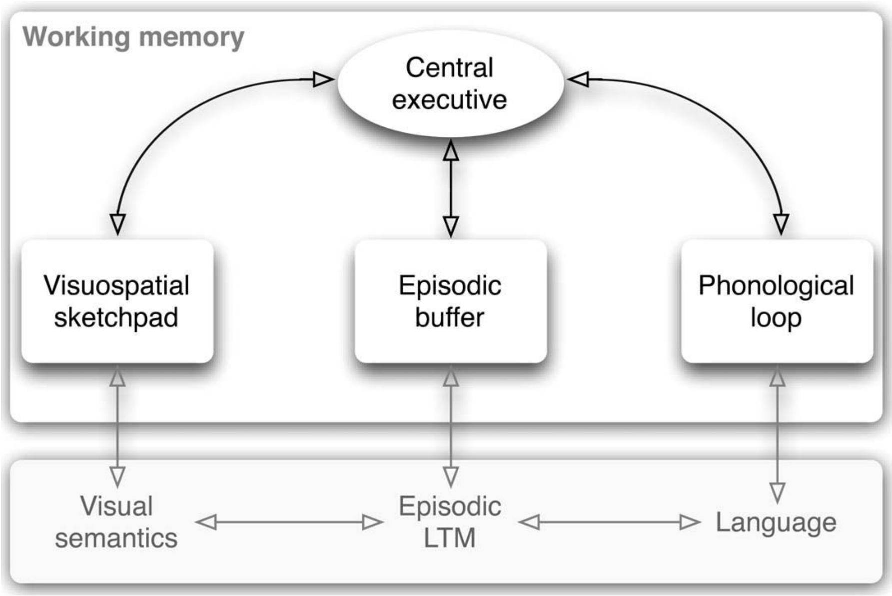


Figure 2 The current multi-component model of working memory representing "fluid" capacities (such as attention and temporary storage) that do not change by learning and their proposed relations to "crystallized" cognitive systems capable of accumulating long-term knowledge.
(Repovs & Baddeley, 2006, Figure 1, p. 6)

Looking closer at the structure of the phonological loop, it was also found to consist of two components: the phonological store and the articulatory rehearsal process. The former stores the memory traces in acoustic or phonologic form. The latter's functions is to retrieve the information in the phonological store and then to rehearse them (Repovs & Baddeley, 2006). The rehearsal process is subject to development: "Before 7 years of age, spontaneous rehearsal does not reliably occur in younger children" (Gathercole et al., 2004, p.177). Thus, in the phonological loop of younger children, there exists just the phonological store.

The central executive is the most important part of the working memory model. However, its functions are less well understood (Baddeley, 1986, 1996, as cited in Repovs & Baddeley, 2006). It is generally assumed to function as the administrator of the other two subsystems. Still, many unanswered questions about its facilities remain. However, in the new model where it is complemented with the episodic buffer it has become clearer on which resources it draws to control the other two slave systems.

Development of working memory in childhood

The ability to store information in memory develops significantly during childhood. This development includes many components like "perceptual analysis, construction and maintenance of a memory trace, retention of order information, rehearsal, retrieval and redintegration[1]" (Gathercole, 1999, p. 410).

The below figure shows the development of the working memory components like phonological short-term memory, visuo-spatial short-term memory and complex working memory for children between 2-16 years of age.

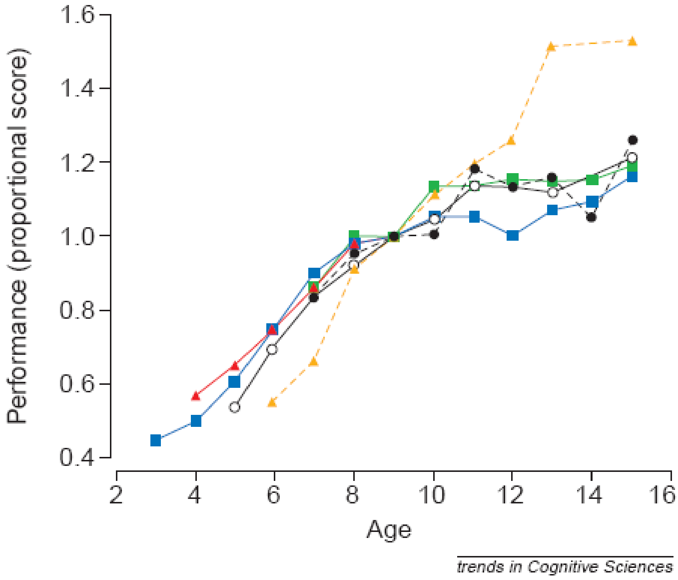


Figure 3 Performance on measures of short-term memory as a function of age. Mean performance of each age group is plotted as a proportion of mean performance of nine-year olds. Blue squares, digit span (phonological memory); red triangles, non-word repetition (phonological memory); open circles, forward digit span; green squares, Corsi blocks (visuospatial memory); yellow triangles, listening span (complex working memory); filled circles, backward digit span (complex working memory). (from Gathercole, 1999, Figure I, p. 411)

Gathercole (1999) explains that the working memory performance increases rapidly up to 8 years of age and then shows a slower improvement until 11-12 years of age.

1 Redintegration is "a reconstruction process in which permanent representations are used to help re-build information from the temporary memory trace" (Brown & Hulme, 1995, 1996; Hulme et al., 1997; Schweickert, 1993, as cited in Thorn et al., 2005, p. 134)

Other than the LST which measures the complex working memory span, all test results are similar to each other. The reason why the complex working memory develops later than both phonological short-term memory and visuo-spatial short-term memory may be related with the lengthy development of the frontal lobes, which presumably support the complex working memory capacity in the brain (Baddeley 2003).

The neurological processes underlying the working memory continue to develop during childhood. Both the audio-spatial working memory and visuo-spatial working memory performance increase with age and the related areas in the brain maturate, as well. There are some developmental differences in favor of girls, especially between the ages 6-10. Boys show more immaturity than girls during this period (Vuontela et al., 2003).

1.2 Categorization

In this study, we also want to investigate the categorization capability of children by means of the categorical free recall experiment. In measuring the release from proactive interference by means of this experiment, we also measure the categorization ability. There should be an interaction between the categorization skill and release from proactive interference because only if subjects do have categories can they show release from proactive interference. Since the release from proactive interference and the categorization capability are strongly connected, we can observe this skill by examining the results of the release from proactive interference task. What we mean by categorization is "to respond differently to objects or events in separate classes or categories" (Ashby and O'Brian, 2005, p.83).

Categorization is an important developmental research issue since this ability can already be observed in very young infants. The reasons why we start to categorize objects already at the very beginning of our life might answer the question of how human beings developed such a complex brain system (Quinn, 2002, p. 66).

Ashby & Maddox (2005) survey several studies that deal with human category learning. Most theories of the category learning assume that there is just one category learning system that allows people to learn all kind of categories. Prototype theory, one of these theories, posits that learning categories is the same as learning category prototypes. When an undefined stimulus is encountered, it is assigned to the most similar category prototype.

Considering the beginning of the categorization ability in human beings, we see that it starts emerging by the second year of life. At this age infants are able to sort objects into categories (Younger and Fearing, 1999). This ability increases throughout their life time. As children get older, they learn many categories and they also acquire the ability of learning new categories quickly. (Ashby and O'Brian, 2005) In a study of Hasher et al. (2002), it was found that release from proactive interference is stronger for adults than for younger subjects. Therefore, in our experiments with adults we expect to see more developed categorization ability than we expect in children. Likewise, in children, we expect to see a stronger categorization effect in the older children.

1.3 Proactive interference

The ability to store information in the brain spontaneously develops in childhood. We become better at following our long-term goals and ignoring the irrelevant information that constrain us in achieving these goals during childhood and adolescence (Bunge & Wright, 2007).

The seminal studies on PI and RPI are quite old already, e.g. from 1945, but the topic has always been studied until recently.

Previously learned behavioral patterns frequently interfere with the learning of new behavioral patterns. These kinds of interferences have been illustrated in the literature even in infant studies such as hidden objects experiments, and the simple rule learning studies with children and especially the theory of mind studies (Carlson & Moses, 2001; Dempster, 1992; Dempster & Corkill, 1999, as cited in Kail, 2002).

In a theory of mind task, for example, the child has to acknowledge that the change in the location of an object has occurred so that her own old representation needs to be updated while the representation of another person might not have been updated since this person did not witness the change. Besides, the ability to inhibit the previously learned information is a key for executive functions (Roberts & Pennington, 1996; Zelazo, Carter, Resnick, & Frye, 1997, as cited in Kail, 2002). Executive functions generally include “the processes of planning and goal-oriented behavior, self regulation, and cognitive flexibility—processes thought to reflect functioning of the prefrontal cortex. Each of these features of executive functioning depends critically on being able to overcome interference from previous experience, including overriding habitual motor responses or ignoring stored information that is no longer needed or accurate (e.g., the temporary location of an object)” (Passler, Isaac, & Hynd, 1985, as cited in Kail, 2002, p. 1703).

Kail (2002) also surveyed the nature of PI effects in children between 4-13 years of age in a meta-analysis including 26 studies and 82 data sets). Many of them report a decrease of PI across age, consistent with the idea of increasing executive functions. However, he continues that also mixed results have been reported in the literature. One account proposed that the strength of PI does not change through childhood (Kail & Levine, 1976; Kail & Schroll, 1974; Tyrrell, Pressman, Cunningham, Steele, & Thaller, 1981, as cited in Kail, 2002). Other studies even found that the amount of PI increases with age (Geis, 1975; Kee & Helfend, 1983; Nakayama & Kee, 1980, as cited in Kail, 2002).

According to Darling & Valentine (2005), when items from the same category are represented, proactive interference (PI) occurs. According to these authors, the items in semantic memory are connected to each other via their common features. Stimulating one item in one category causes the other items to become activated too (priming). Thus, the other items in the same category can be activated easily. However, the learning of previous items interferes with learning new items, in particular of the same semantic category. This is the effect of PI. If, however, members of a new semantic category are encountered, release from proactive interference (RPI) is observed, that is, this new item does not suffer from PI anymore but is remembered much better.

The studies in the release from proactive interference area rely on the following 4 findings (Derwing & de Almeida, 2004):

1. Short-term memory has several constraints like the magic number 7 +/- 2 (Miller, 1956).
2. Generally, recall of information decreases over a brief period of time, specifically, items in short-term memory will decay if they are not rehearsed in the phonological loop (Peterson & Peterson 1959, as cited in Derwing & de Almeida, 2004).
3. The storage of information in memory is prone to proactive inhibition (i.e., PI).

In the PI task, originally developed by Wickens (1970, 1972, as cited in Derwing & de Almeida, 2004), participants receive 3 sets of items followed by the interference task. After this interference task, they are required to remember the items at the beginning of the experiment. All these, namely three sets of items, the interference task, and the recall, are repeated three times. According to the results, if all items are from the same category, recall accuracy drops significantly towards the end of the list which exemplifies proactive interference. However, if the items in the interference task are from a different category, recall accuracy is as good as in the initial set which exemplifies release from proactive interference.

Generally, the studies on release from proactive interference include several (procedural) differences relating to the stimuli and the representation of the stimuli. Therefore, the results in this area are quite divergent. For example, the three studies which examined the ability to use the evaluative dimension in the encoding process by means of the RPI method for children differed in their results (Cermak et al., 1972; Kail & Schroll 1974; Pender, 1969, as cited in Douglas & Corsale, 1977). In these studies, children were required to recall 4 sets of positive words (e.g. fresh, open, nice) including 2 items each. Then, in the 5th set negative words (e.g. worry, burn, hate) were used or the vice versa. A color-naming task was used between the sets as a distractor. The results were mixed. While the first two studies observed the

RPI effect both for the 2nd and the 6th graders, in the last study the RPI was not present for the 2nd graders.

Gardiner, Craik, and Birtwistle (1972, as cited in Darling & Valentine, 2005) have investigated proactive interference and retroactive interference paradigms. They used both categories (e.g., flowers and games) and sub-categories (garden flowers vs. wild flowers) in their experiments. As a result, while release from proactive interference could be observed between the superior categories as well as for the sub-categories, RPI was only observed if the participant had been informed about the specific category types previously.

Also, Marques (2000, as cited in Darling & Valentine, 2005) used the release from proactive interference technique in order to examine the difference between living/non-living objects in semantic memory, with pictures as well as words. In his experiments, he used a category shift between living and non-living objects. The study showed an RPI effect for words but not for pictures.

Moreover, Wickens (1970, 1972, as cited in Douglas & Corsale, 1977) adapted the Peterson and Peterson task (1959, as cited in Douglas & Corsale, 1977) for adults. In this experiment, participants receive 3 sets of items from the same category. Then, the items in 4th set changed into a different category or remained the same. Since all items were from the same category, there would be a build-up of proactive interference. When the 4th set contained items from the same category recall performance dropped significantly. However, when the last category changed, recall performance reached the initial level.

In order to explain the RPI effect, Winston (1999) discusses three different hypotheses. The first one is the *attentional* or *encoding hypothesis* which maintains that participants be alarmed about the change in the list so the processing mechanism of the participant increases which leads to the item's better storage in memory (Wickens, 1970, as cited in Winston, 1999). The second hypothesis is the *storage hypothesis*. It claims that there is an interaction between the current item and previously stored similar items during the process of proactive interference. It proposes the metaphor that memory is like a library in which items are stored on shelves of their own and that similar items are closer to each other in this library, i.e., they are on the same shelf. Thus, a shift in the list will be less affected by the interference since the new items are not any longer similar to the previous items in memory. The last hypothesis is the *retrieval hypothesis*. It maintains that each item in memory is encoded equally but that the retrieval process gets difficult because of the increasing effect of proactive interference. Therefore, changing the type of item increases its efficient retrieval. Thus, a shift in the material would provide more powerful retrieval cues for the participant.

Engle and Kane (2003, as cited in Hamilton & Martin, 2007) report that susceptibility to interference has been found to be related to working memory span because low span participants display more interference effects than high span participants. The authors propose that "the executive control of attention is responsible for the relationship between working memory and interference resolution" (p. 113).

Hasher et al. (2002) point out that in RPI tasks, consecutive lists are recalled which then immediately become irrelevant to the subject. If the deletion is successful, the previous items will be suppressed and this will make current items easy to remember (Underwood, 1957, as stated in Hasher et al., 2002). If the deletion is not successful, there will be poorer recall. The deletion requires executive functions since singular items or whole lists of items have to be actively suppressed. This deletion ability is subject to development and ageing (see section 2.4)

Lastly, it has been claimed that the RPI effect might increase because of the linguistic characteristics of the items in the task, such as semantic attributes, word length, and phonological similarity (Winston, 1999).

1.4 Neuro-Anatomical Basis of PI and ageing

The proactive interference paradigm has been investigated in several behavioral studies on both short-term memory and long-term memory. Just and Carpenter (1999) argue for its relevance in those areas as follows: "As a determinant of working memory capacity, it is also important to higher cognitive functions that depend on working memory" (p. 107, as cited in Du et al., 2008).

Anderson and Neely (1996, as cited in Jonides &. Nee, 2006) state that it is actually proactive interference, resulting from the competition for retrieval, which seems to diminish working memory capacity. Namely, the working memory capacity reflects the number of items to be stored in memory and proactive interference makes it harder to retrieve information from memory.

In proactive interference, a past experience interferes with a subsequent experience. In order to overcome this interference, executive functions are needed. They are generally supported by frontal regions of the brain (Baddeley 2003). More specifically, some of the relevant studies (as discussed in Badre & Wagner, 2005) imply "the left mid-ventrolateral prefrontal cortex (mid-VLPFC) in PI resolution during short-term item recognition" (p. 2003). However, the mechanism is not well-understood yet.

The perception of the present is assumed to be shaped by our past experiences and some of them might be detrimental as in the process of proactive interference. Proactive interference has been considered as a distractor for memory and cognition and also as a reason for long-term forgetting (McGeoch, 1942, as cited in Badre & Wagner, 2005). It has also been shown to be responsible for age-related decline in cognitive skills (Hasher and Zacks, 1988). Proactive interference can limit the memory processes and thus lead to short-term forgetting, as well (Brown, 1958; Peterson and Peterson, 1959; Keppel and Underwood, 1962, as cited in Badre & Wagner, 2005). Thus, "the processes that resolve or resist PI may be critical for the flexible updating and maintenance of task-relevant goals, stimuli and responses." (p. 2003).

Since proactive interference has several detrimental effects on memory, its resolution is important for goal-relevant behavior. Badre & Wagner (2005) found that various cognitive control mechanisms helped resolve PI, some of which were supported by left mid-VLPFC and PFC.

Jonides and Nee (2006) argue for the necessity to study release from proactive interference and its brain bases. In the case of increased retrieval competition through proactive interference, context-retrieval may help to resolve the competition. Context-retrieval is mediated by the left inferior gyrus. Given that working memory is an important predictor of other cognitive skills and proactive interference limits its capacity, it is worthwhile to further study mechanisms that may resolve proactive interference.

In the same vein, Nee et al. (2007) critically emphasize the need to diminish proactive interference for the sake of successful cognitive processing. Left ventrolateral prefrontal cortex (VLPFC) was found to be related with this ability in a number of studies. However, the contribution of VLPFC could not yet be generalized for different tasks. This study, using event-related functional magnetic resonance imaging (fMRI), found that "both left VLPFC and left anterior prefrontal cortex (APFC) are involved in the resolution of proactive interference across tasks" (p. 740).

1.5 Assessing RPI and related WM abilities in experiments

Proactive interference is related to executive functions of working memory as well as to phonological working memory, insofar as lists of words have to be learned. Therefore, we will not only administer a PI experiment but also additional measures that are related to working memory, two for the functions of executive/complex working memory, and one for phonological working memory. The WCST and the LST (for children) as well as the Reading Span Test (for adults) are our measures of executive and complex working memory. The word-span test is for measuring phonological working memory. These tests are classical ones to measure these functions. All these tasks are therefore considered valid and reliable.

The main aim of carrying out these additional experiments is to find out what kinds of functions of the human working memory system are related with the release from proactive interference memory task.

In the following, some background information is given on the various kinds of memory tasks related to RPI, executive functions and phonological working memory, as they are discussed in the literature. More specific information about our own tasks (stimuli, procedure, analysis), will be given in Chapter 3 on "Methods".

1.5.1 Categorical Free Recall Test (CFR-Test)

Release from proactive interference experiments are generally conducted in the following way: There are three trials with items belonging to the same category which is then followed by a fourth trial which is from another category (Halford et al., 1994). The first 3-trial set causes the building-up of proactive interference and the category shift (between the 3rd and 4th item) causes the release from proactive interference.

Kee et al. (2005) also examined the RPI mechanism by developing a gender schema encoding for occupations. It is the only article we found about the RPI effect in children. This study is relevant to the current study in terms of the procedure/method they used and on the developmental issue but not so much on the content of their categories, namely gender-based occupations and their interest in the activation of gender stereotypes. In their experiments, they used some typically feminine (e.g., ballet dancer, nurse, secretary) and masculine (e.g., police officer, electrician, and dentist) occupations. Three sets of items including 4 items from feminine occupations were used and then the items in the 4th set were from the masculine occupations, or the vice versa. Figure 4 shows that in the adult sample, RPI occurs in the experimental condition (where the gender is changed after the third set) but not in the control group (where the gender is maintained). It seemed that RPI only occurs for young adults but not for children (6-graders, 11 yrs).

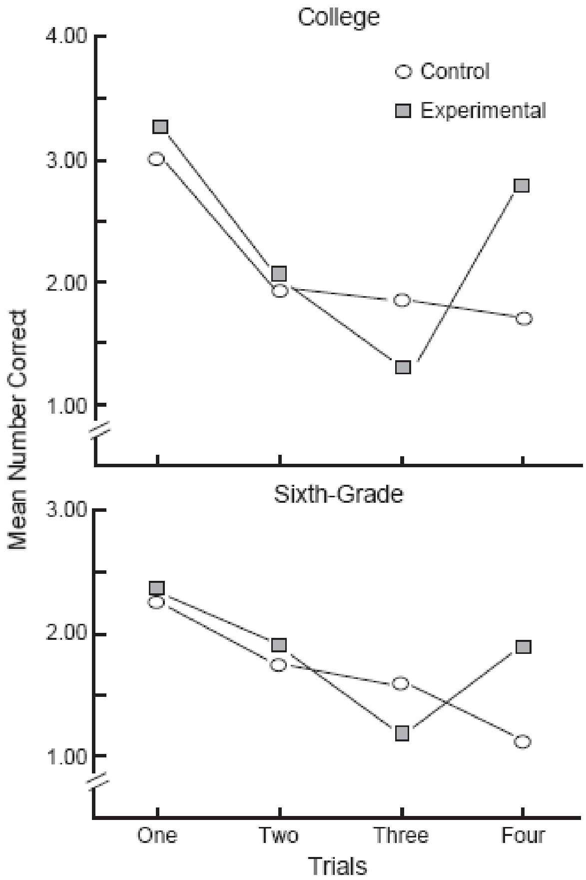


Figure 4 Mean number of correctly remembered items as a function of condition and trials for college (upper panel) and sixth-grade (lower panel) participants (Kee et al., 2005, Figure 1, p. 207)

In the current study we will be concerned with much more basic semantic categories across a longer developmental period (grade 1-5 and adults). Our own RPI task will be similar in structure; however, it will be composed of three categories with four items in each. The overall 12 items have to be recalled freely (in correct serial order). Therefore we call it "Categorical Free Recall Test (CFR-Test)".

1.5.2 Word Span Test (WST)

In the Word Span Test (WST) (Pickering & Gathercole, 2001, as cited in Alloway et al., 2004) children are required to recall a given set of words in the presented order. At each level, the number of items increases by one until the child makes an error on one of the 4 trials at each level. Then, the word span equals the correct number of words the child can remember. The original WST was adapted from English to Turkish.

1.5.3 Wisconsin Card Sorting Test (WCST)

Firstly, the WCST can be defined as a complex memory measure (Cianchetti et al., 2007) since the WCST performance is included in working memory which was assumed to be a fundamental part of executive functions (Stratta et al., 1997).

As Stratta et al. (1997) claim the WCST is the most used neuropsychological test in the literature. It is “a complex task involving learning, elaboration of strategies for hypothesis testing and problem-solving” (p.18). The WCST is also used for children. In this case it is not much a neuropsychological test but a test for the development of executive functions.

1.5.4 Listening Span Test (LST)

Gathercole & Alloway (2004) suggested that the working memory capacity in children can generally be measured with complex span experiments in which both storing and processing of information are required.

Experiments on verbal short-term memory and visuospatial short-term memory require only storing of information in memory while working memory experiments comprise both processing and storage of the spontaneous memory, like in the listening recall test (Alloway et al., 2006).

In the LST, children are required to listen to some simple sentences like “Apples are red” and then decide about the truthfulness of this sentence. At the end of each sentence sequence (whereby the number of sentences increased step by step), the last word of each sentence has to be recalled. The experiment begins with one sentence and increases one by one until the child is unable to recall the three last words for the three sentences in one sequence (Archibald & Alloway, 2008). This test was also adapted from English to Turkish.

1.5.5 Reading Span Test (RST)

As Gathercole et al. (2004) state the RST is one of the tests used to investigate the Central Executive in adults. Whitney et al. (2001) explained that in the RST, participants are required to read some sentences aloud and then to judge the truthfulness of the sentences. At the end of each set of sentences, they are required to recall the last word of each sentence. There are 3 sentences in each set of sentences. The set size varies between 2 and 6. The test results can be evaluated in two ways: Either the largest set size that the participant completed correctly or the total number of remembered target words could be taken as the score (Daneman and Carpenter, 1980 as cited in Whitney et al., 2001).

In Saito & Miyake's (2004) RST, the target words were either at the end of the sentence or in near-final position. (In the original English RST, the target words were always in final position.) However, since in Japanese (the language in which Saito & Miyake carried out their study) in the final position words are usually verbs, they changed the original test by also allowing target words in other than final position. We followed this reasoning for our Turkish adaptation of the RST. Another reason for not always using the final position is that it becomes easier to predict the position of the target words.

1.6 Socio-economic effects related with the planned tests

Socio-economic factors on (working) memory performance are well discussed in the literature. Generally, the socio-economic factors taken into consideration are birth order, age spacing between siblings, socio-economic level of the family, and family size. The studies that dealt with the psychological characteristics of individuals including birth-order of the subject found that socio-economic factors were indeed very critical, however, other factors such as birth order, were not that important as originally suggested (Schooler, 1973, as cited in Cicirelli, 1978).

Firstly, Bayley (1965, as cited in Cicirelli, 1978) proposed that until 15 months of age birth order or gender effects would not influence intelligence scores. Then, Anastasi (1956, as cited in Cicirelli, 1978) found a negative relationship between family size and intelligence. However, she continued that this effect disappeared when the socio-economic level increased. Also, Glass, Neulinger, & Brim (1974, as cited in Cicirelli, 1978) pointed out that in families from higher socio-economic levels first-born children had better scores than later-born ones. However, no significant effects in this respect were found in families from the lower socio-economic level.

Two hypotheses are discussed for the family size effect: The *dilution hypothesis* says that since the attentional resources are limited within a family, the more children there are in a family the less attention is given to each child individually (Blake, 1981; Downey, 2001, as cited in Bjerkedal et al., 2007). The *confluence model* says that a child is affected by the overall intellectual capacity in the family and also elder siblings have the tutoring function (Zajonc, 1976, 2001a, p. 513, as cited in Bjerkedal et al., 2007).

Cicirelli (1978) asserts that the age spacing between the siblings is also another important supportive factor. If the age spacing is longer for two siblings, then the older child would have well-developed intellectual abilities and for the smaller one the average intellectual level of the family would be better. Every time a child enters the family, the average intellectual level would drop. Also, as Breland (1974, as cited in Cicirelli, 1978) claimed that if the age spacing between the two siblings is lower than 2 years their intellectual scores would be quite low.

Boomsma et al. (2008) indicate that there is a negative relationship between the IQ and the birth order. There are also conflicting outcomes for the same relation, however. The reason for these contradictory findings might be the age on which IQ is measured (Sulloway, 2007, as cited in Boomsma et al., 2008).

Generally, no differences were found between the first-born child and the later-born ones. Thus the previous studies which had found opposite results might have ignored some other factors such as birth order, family size, age spacing, etc. (Cicirelli, 1977, as cited in Cicirelli, 1978).

Also, some studies report that it is very important to consider the cultural structures of the families since the relationship between the siblings and the importance of the birth order is affected by the culture of the society (Cicirelli, 1978) so that these factors can have a different valence in different cultures.

In order to address the main socio-economical variables that are known to have an impact on the development of intelligence and working memory, we devised a questionnaire in which we asked for the relevant information (see also Appendix B).

CHAPTER 2

METHOD

2.1 Participants

All experiments were carried out with both children and adults as a control group. Children's age range was between 6-12 years. They came from grades 1-5. The overall sample consisted of n=101 children. Both the descriptive statistics and the age values are shown in Table 1 and Table 2. The pupils were recruited from two primary schools in a rural area in Yozgat, in the Central Anatolia Region of Turkey.

In the adult group, there were 20 graduate students from the Informatics Institute, METU, especially from the department of Cognitive Science. The age range of the adults was between 23 and 30 years.

Table 1 Descriptive statistics for the children group

	N	Minimum	Maximum	Mean	Std. Deviation
Age of the subject in years	101	5,58	12,08	8,5520	1,44900

Table 2 Age values for each class

	N	Mean	S.D.
Grade 1	20	6, 711	,125
Grade 2	24	7, 591	,092
Grade 3	16	8,691	,178
Grade 4	22	9,504	,087
Grade 5	19	10,486	,143

Table 3 Descriptive statistics for the adult group

	N	Minimum	Maximum	Mean	Std. Deviation
Age of the subject in years	20	22,34	30,89	25,8520	2,04300

2.2 Experiments

2.2.1 Categorical Free Recall Test (CFR-Test)

Material

There were 3 different categories (fruits-animals-clothes) and 12 words in the test (for further details see Chapter 0):

Fruits : erik, karpuz, üzüm, çilek (plum, water melon, grape, strawberry)

Animals : tavuk, aslan, maymun, inek (chicken, lion, monkey, cow)

Clothes : kocuk, çorap, kazak, gömlek (coat, sock, sweater, shirt)

Procedure

During the test, subjects listened to the items from the above list via a headphone. There were 18 different orders of items and categories. Immediately after having listened to the items, the subjects were required to recall what they remembered from the list in serial order. The answers of the subjects were recorded by a voice recorder (See Appendix B for the ethical details of the study).

2.2.2 Word Span Test (WST)

Material

The WST comprised one-syllabic words from Turkish since the main task consisted of 2-syllabic words. Thus, we do not want these two tests resemble each other. Still, it might be objected that the recall of one-syllabic words may actually be harder than the recall of two-syllabic words, in terms of overall number of syllables retained. However, using one-syllabic words is the most conservative and therefore most basic measure of their memory span. The words in this task have been chosen with respect to easy pronunciation and high frequency in daily usage, such as "saç, tuz, türk, and yurt" (hair, salt, Turkish/Turk, and country). We constructed various sets by using these words. The longest set size is 8, and the smallest set size is 2. Also, on each level of set size there are three sets in the experiment. An example of a 2-word-set is the following (see Appendix A for the whole stimuli list):

1. top	can	(ball	soul)
2. bil	kürk	(to know	fur)
3. ver	tez	(to give	quick)

Procedure

In the experiment, the subjects listened to the sets of word with increasing length. The sets began with the smallest set size, i.e. 2, and then increased with the performance of the subject. If the subject made two or more mistakes in an overall set, the experiment was terminated. The answers of the subjects were recorded via a voice recorder.

2.2.3 Wisconsin Card Sorting Test (WCST)

Material

In the basic WCST, there are two sets of 64 cards. However, in the Modified WCST for children (Cianchetti et al., 2007), there are two sets of only 24 cards. These cards show different combinations of colors (yellow, red, green, blue), numbers (one, two, three, four), and forms (triangle, star, circle, cross).

Procedure

In the WCST, there were four stimulus cards (see Appendix A) which were shown to the subjects during the experiment. These stimulus cards were a single red triangle, two green stars, three yellow crosses, and four blue circles, respectively. The remaining response cards had all the combinations of different colors, forms, and numbers (Cianchetti et al., 2007). The response cards were displayed one by one to the subjects. Then, for each response card, the subject was required to give an answer by indicating a stimulus card for each response card. The answers of the subjects were written on the answer sheet for that subject.

In the normal WCST for the adults, the rules were as follows (Cianchetti et al., 2007):

- The required number of consecutive correct answers was 10.
- The order of categories in the experiment was: color-form-number-color-form-number.
- After 10 consecutive correct answers, the sorting criterion was switched; however, the participant was not informed that s/he has to find another rule.

However, in the Modified WCST (Cianchetti et al., 2007) for children, the rules were changed in the following ways:

- 2 sets of 24 cards instead of 2 sets of 64 cards.
- The required number of consecutive correct answers was 6 (10 in WCST).
- Cards which shared more than one attribute (color, shape and number) were excluded.
- Whatever category the participant chose first was taken as correct.
- After 6 consecutive correct answers, the participant was informed that s/he had to find another rule.
- After the three categories (color, form, and number) were completed, the remaining 3 categories were the same as the previous ones.

2.2.4 Listening and Reading Span Test (LST & RST)

Material

LST

In the LST for children, there were sets of sentences belonging to the some bigger sets in the experiment. The longest set size comprised 6 sentences and the smallest set size 2. On each level of set size, there were 6 sets of sentences. For example, on the 2-sentence-set level, there were 6 groups of 2 sentences. In the experiment, the total number of sentences was 212. An example of a 3-sentence set for the LST is the following one (see Appendix A for the whole stimuli list):

1. Muzlar bisiklete biner. (Bananas ride bicycles)
2. Elimiz beş parmaklıdır. (Our hands have five fingers)
3. Soğan acıdır. (Onions are hot)

RST

The Turkish Reading Span Test for the adults was created by the author. The test is similar to that of Saito and Miyake (2004). In this experiment, sentences were selected from school books for children of medium age consisting of widely known facts. The structure of the test was the same as that of the LST. An example of a 3-sentence set for the Turkish Reading Span Test is the following one (see Appendix A for the whole stimuli list):

1. Salon sporlarından biri de bowlingdir.
   (One of the salon sports is bowling)
2. Sebzeler bol miktarda B vitamini ihtiva eder.
   (There is much vitamin B in vegetables)
3. Osmanlı Devleti dünyadaki en uzun süren imparatorluktur.
   (The Ottoman Empire is the longest lasting empire in the world)

Procedure

LST

In the experiment, the subjects listened to the sentences and then were required to answer “Yes“or “No” according to the best of their knowledge. In addition, they also had to keep in mind the last word of each sentence. For example, when they listened to the sentence “İnekler uçar” (Cows can fly), then they should say “Hayır, uçar”.

The set size of the test increased with the performance of the subject. Like in the WST, if the subject made two or more mistakes in one sentence set, the experiment was terminated. Also, the answers of the subjects were recorded by a voice recorder.

RST

This experiment was very similar to the LST. The first difference was that, here, the subjects read aloud the sentences from the screen. They then indicated whether the sentence was true or not according to the best of their knowledge. The second difference was that the subject should remember not the last word of the sentence but the underlined, red, target word of the sentence. At the end of each sentence set, the subject were required to recall what he remembered as target words in this set. Also, the cut-off criterion was the same as in the LST.

2.3 Analyses of the tests

After all experiments had been carried out, I did a statistical data analysis on all of these tests (the CFR-Test, the WST, the WCST, the LST, and the RST). As the main statistical procedure variants of ANOVA were used. For each experiment, I included age as a factor in the ANOVA. In order to investigate the relation between the tests and the proactive interference task I ran multiple regression analyses. My predictors in the regression were LST/RST/WCST/WST, and the predicted value was the performance in the release from proactive interference task. Since the data was non-normally distributed according to Kolmogorov-Smirnov Test and Shapiro-Wilk Test ($p < .05$), Kruskal-Wallis Test, Wilcoxon Signed Rank Test, Mann-Whitney Test, was used as the non-parametric tests. Since non-parametric tests did not allow including more than one variable, ANOVA was used in order to examine the interactions between gender and grade.

CHAPTER 3

RESEARCH QUESTIONS & HYPOTHESES

My general research questions are the following ones:

1. How does release from proactive interference (RPI) as measured in the CFR-Test task develop in Turkish school children (age 6-12 years)?
2. How does the development in this memory task relate to other cognitive capabilities that the child has to have in order to solve this task (relations with executive functions, phonological and complex working memory)?

These developmental research questions will be addressed on the background of mature adult performance. We therefore conducted all experiments with adults as a control group, too.

The hypotheses of the study are the following ones:

- **H1**: Overall memory span increases with age. Consequently, the number of remembered items should increase with the age of the children in the CFR-Test.
- **H2**: The categorization capability should increase with the age of the children in the CFR-Test, that is, we expect to see PI across items within the same category and RPI between categories. This capability should be clearly pronounced in the older children in particular.

- **H3**: With respect to proactive interference we have no directed hypothesis (as to whether it should decrease or increase over age, see literature review in Chapter 2). However, the amount of PI will obviously interact with the RPI in that the level of recall reached at the end of a category will partly determine the RPI effect. However, the RPI should increase with age due to the greater categorization ability of the children that helps them realize the beginning of the new category more clearly.
- **H4**: The performance of the adults should be higher than the performance of the children both in the number of remembered items and the categorization capability, that is, the characteristic pattern of build-up of PI and subsequent release should be more pronounced.
- **H5**: The additional memory tasks – the Word Span Test (WST), the Wisconsin Card Sorting Test (WCST), and Listening Span Test (LST) – should also show developmental progression.
- **H6**: These additional tasks should predict the results from the main task to a significant extent since all tasks underlie similar mechanisms in short term/working memory. In particular the executive functions (as measured by the Wisconsin Card Sorting Test (WCST) and the Listening Span Test (LST)) should be predictive of the results of the main tasks.

CHAPTER 4

RESULTS (CHILDREN)

4.1 The CFR-Test

4.1.1 Overall memory performance

For the CFR-Test, the number of subjects, the mean values and standard deviations are shown in Table 4. Absolute order of items (&lists) calculates the number of correctly remembered words according to the serial order of the items (&lists) in the list. Relative order of items (&lists) calculates the number of correctly remembered words according to the order of the item in the list but irrespective of the categorical order. Overall, from a 12-word list, children remembered 5, 56 words.

Table 4 Descriptive statistics for the main task

CFR-Test	Mean	S.D.	Median
absolute order of items	1,08	1,146	1,00
absolute order of items&lists	2,35	1,808	2,00
relative order of items	3,04	1,939	3,00
relative order of items&lists	4,48	2,360	5,00
total # of recalled items	5,56	1,396	6,00
total # of recalled items&lists	8,38	1,766	9,00

We tested the significance of these above values with respect to grade. Since the data were non-normally distributed according to the Kolmogorov-Smirnov and Shapiro-Wilk Test (p < .05), the Kruskal-Wallis Test was used. According to the results, grade has a significant effect on absolute order of remembered items, absolute order of remembered items & lists, the number of remembered items, and the number of remembered items & lists ($\chi^2(4) = 9.49$, $p = .05$, $\chi^2(4) = 12.06$, $p < .05$, $\chi2(4) = 9.59$, $p < .05$, and $\chi2(4) = 9.76$, $p < .05$).

The working memory capacity develops both in early and middle childhood (Gathercole & Alloway, 2004). This finding was supported by our data. The overall results of this experiment were in line with the hypothesis that the number of remembered items should increase with the age of the children in the CFR-Test. Figures 5 and Figure 7 show the development of the overall number of remembered items and the absolute order of items and lists across the five grades, respectively.

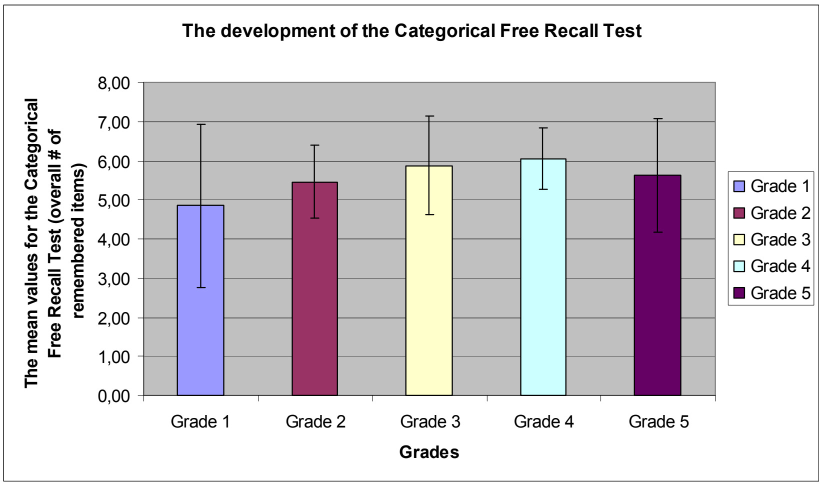


Figure 5 The development of the CFR-Test for the overall number of remembered items (Error bars represent SDs)

For the main task, we found no significant differences between the boys' and the girls' results according to the Mann-Whitney Test's results (absolute order of items, Z = -.720, p = .471, absolute order of items&lists, Z = -.386, p = .699, relative order of items, Z = -.577, p = .564, relative order of items&lists, Z = -.392, p = .695, total #of recalled items, Z = -.976, p = .329, total #of recalled items&lists, Z = -.803, p = .422).

Figure 6 represents the gender difference between girls and boys for all variables for the CFR-Test (the total #of recalled items). In the first and second classes, girls' scores were better than the boys' a little. However, in the third class and beyond, the boys' scores became better.

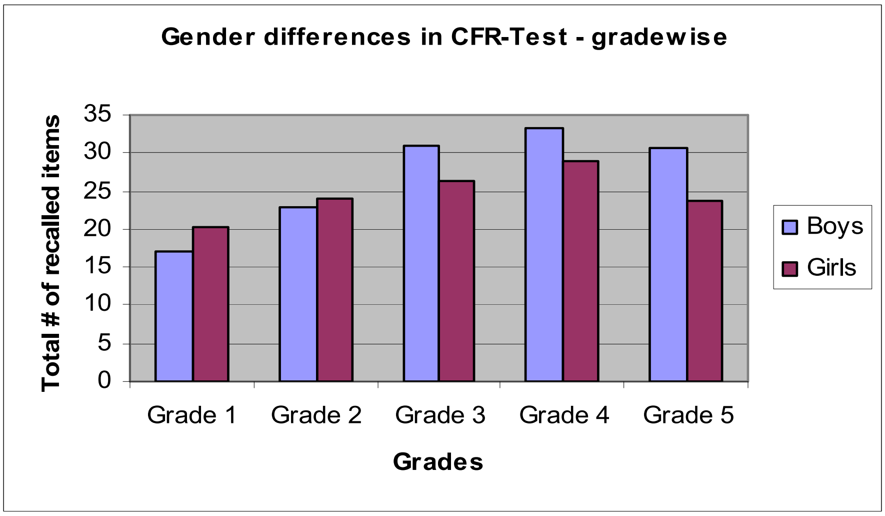


Figure 6 The development of the CFR-Test for the total # of recalled items for girls and boys separately, for each grade

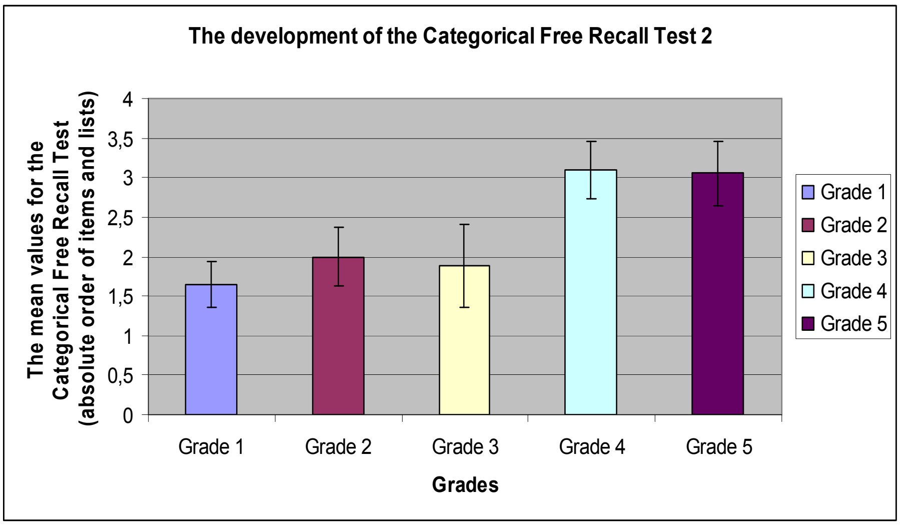


Figure 7 The development of the CFR-Test for the absolute order of items and lists (Error bars represent SDs)

For the absolute order of items factor, boys' results indicated that during the 1st, the 2nd, and the 3rd classes, they got the higher scores. Then, in the 4th and the 5th classes, girls' scores turned out to be better.

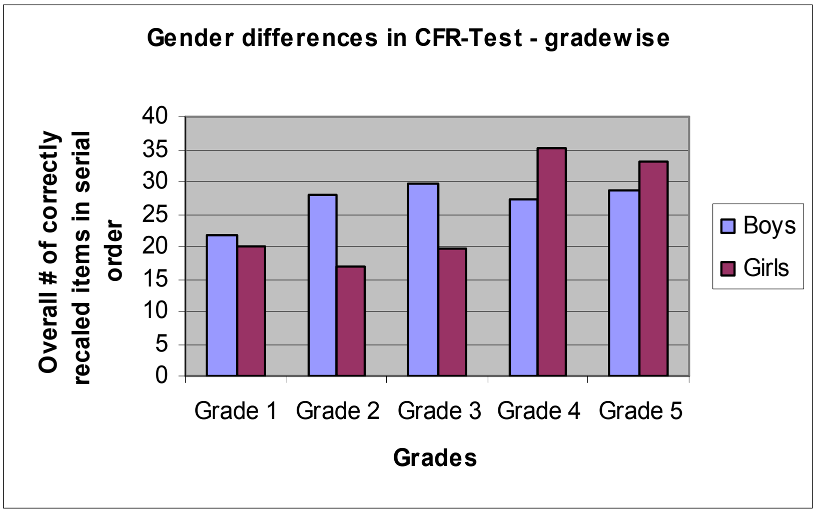


Figure 8 The development of the CFR-Test for the absolute order of items for girls and boys separately, for each grade

We also separated the results of all tests according to the independent variable gender and the grouping variable was grade. For boys, only the overall number of remembered items was marginally significant ($\chi^2(4) = 8.774$, $p = .067$). However, for girls, the results for absolute order of items and absolute order of items & lists were highly significant ($\chi^2(4) = 15.510$, $p = .004$ and $\chi^2(4) = 15.918$, $p = .003$). Thus, boys only marginally develop over time for this sample.

Table 5 The mean ranks for the CFR-Test for the boys

	Grade	N	Mean Rank
Category Test – total # of recalled items	1	10	17,10
	2	12	22,71
	3	8	31,00
	4	13	33,27
	5	10	30,70
	Total	53	

Table 6 The mean ranks for the CFR-Test for the girls

	Grade	N	Mean Rank
Category Test - absolute order of items	1	10	20,10
	2	12	16,92
	3	8	19,69
	4	9	35,33
	5	9	32,94
	Total	48	
Category Test - absolute order of items&lists	1	10	18,70
	2	12	18,17
	3	8	18,38
	4	9	36,28
	5	9	33,06
	Total	48	

We can conclude that in boys it is the capacity of the working memory that (marginally) increases but for the girls it is the serial (and relative) order that develops. When we examine the plots, we see that the increase take places between the 3^{rd} and the 4^{th} classes. However, for the boys it is between the 2^{nd} and the 3^{rd} classes. Although Gathercole et al. (2004) found that there was no significant difference between girls and boys regarding the verbal-storage tasks, for example digit recall, word list recall, and non-word list recall, it seems that for boys the increase in working memory capacity is important while for the girls the development of serial order is important.

Lastly, in section 4.3, it was found that for the CFR-Test, the absolute order of items can be predicted mostly by the LST. It, thus, indicates that there is strong relationship between the executive functions and the ability to recall items in serial order. Since, for the absolute order of items factor, only the girls develop, as Vuontela et al. (2003, p. 74)'s claim is supported that "The gender differences found in the performance of working memory tasks suggest a larger degree of immaturity in boys than girls at the age period of 6–10 yr."

4.1.2 Release from proactive interference (RPI)

In order to assess the results for the release from proactive interference, we ran a Wilcoxon Signed Rank Test as a non-parametric test since the data were not normally distributed (the results of Kolmogorov-Smirnov Test and Shapiro-Wilk Test were $p < .05$). In the overall item set, there occurred two category switches, the first one between items 4 and 5 and the second one between items 9 and 10. These shifts between the last item of the first category and the first item of the second category and the last item of the second category and the first item of the last category are critical for the evaluation of the RPI effect. The following statistical tests were run under a directed hypothesis, namely that the first item of the second category would be remembered better than the last item of the first category. Therefore, we report one-tailed p-values.

Overall, across all grades, for the first category switch, there was a significant change in the number of remembered items between the last item of the first category and the first item of the second category. ($Z = -2.023$, $p = .022$). Also, for the 2^{nd} category switch, for the last item of the second category and the first item of the last category, it was significant ($Z = -1{,}697$, $p = .045$).

More specifically, according to the results of the paired sample t-tests, for grade 1, the 1^{st} difference was marginally significant ($Z = -1.633$, $p = .052$) the 2^{nd} one was significant ($Z = -2.121$, $p = .015$). For grade 2, the 1^{st} difference was marginally significant ($Z = -1.587$, $p = .055$). For grade 3, none of the differences were significant. For grade 4, the 2^{nd} difference was also significant ($Z = -1.897$, $p = .028$). Lastly, for grade 5, the RPI effect disappears again, i.e. the only (marginally) significant difference was the 1^{st} shift ($Z = -1.387$, $p = .086$). To sum up, it can be inferred that release from proactive interference does not increase over age. Release from proactive interference was present in grade 1. It vanishes, however, completely for grades 3 and 5. Otherwise, in all grades, at least for one difference, the difference was significant or marginally significant. Therefore, we conjecture that there might be some other factors responsible for this decrease in release from proactive interference (see section 4.1.2.1).

All figures representing the frequency of items 1-12 in percent for the CFR-Test are shown in the figures below (see Figure 9-13). In these graphs, for each class, all remembered items were summed for each separate item. The bars represent the percentage of children (in the respective grade) that remember this item (irrespective of correct serial order). In other words, the following plots are position-related.

To begin with, in grade 1, children did not show the build-up of PI effect in the first category (see Figure 9). Surprisingly, through the end of this category, the items were remembered better than the former ones. Therefore, the build-up of PI did not show up for the 1st category. However, for the second category and the third category the build-up of PI could be seen easily. Also, both category shifts (between item 4&5 and 8&9) clearly indicated the RPI. Also, the overall amount remembered items was very low.

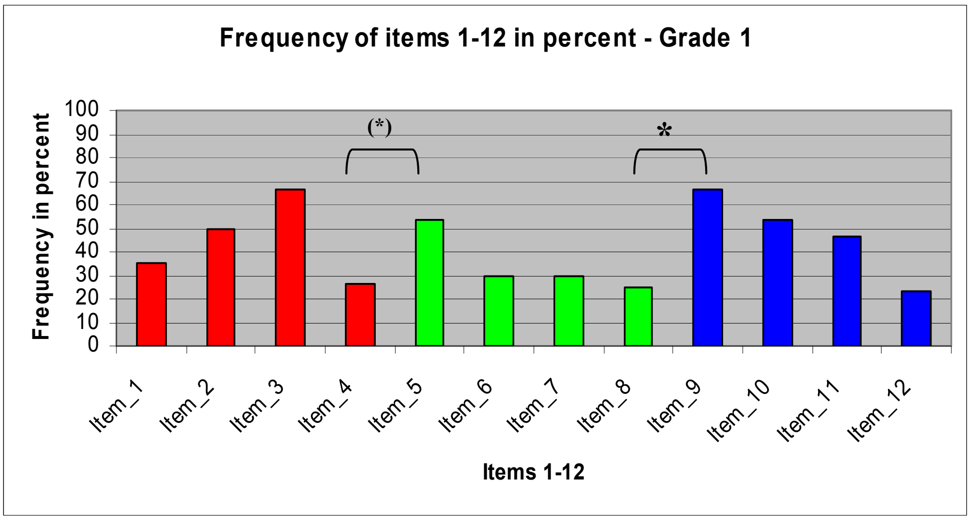


Figure 9 The frequency of items 1-12 in percent for the main task for the 1st class.

In this and in the following figures, an asterisk (with parenthesis) indicates a (marginally) significant difference (hence, RPI), "n.s." marks an insignificant difference.

For the 2nd graders, the pattern mostly changed (see Figure 10). Except the item 1, the first category reflected the build-up of PI. Again, other than the 8th item, the build-up of PI could be seen for the second category. On the contrary, in the last category the levels were nearly the same so there was no build-up. Therefore, just for the first category there was RPI. The absolute level of the memory did also increase for this class.

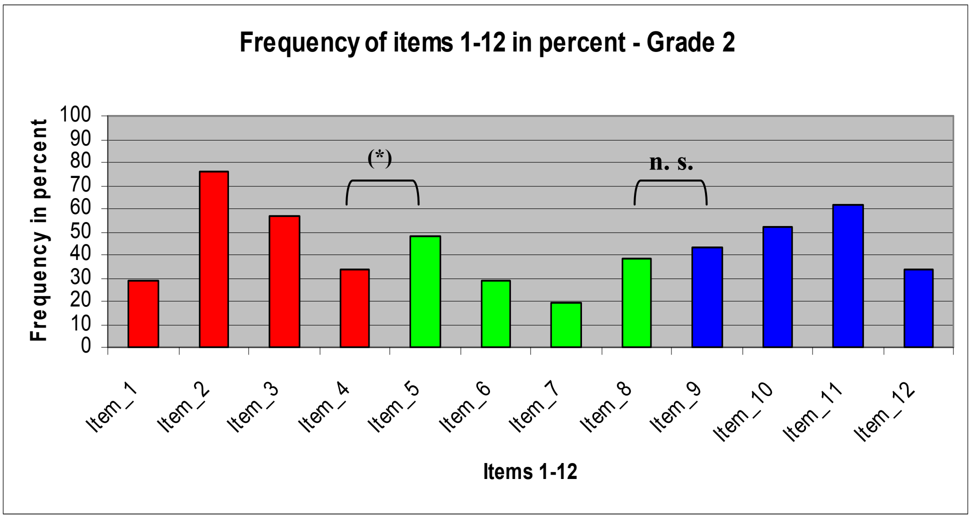


Figure 10 The frequency of items 1-12 in percent for the main task for the 2nd class

The 3rd graders again showed a different pattern. Only for the first category, build-up of PI was present. So the RPI was to be found only between the first and second category. In general, the position curve looks rather flat. This may stem from an overall increase in WM.

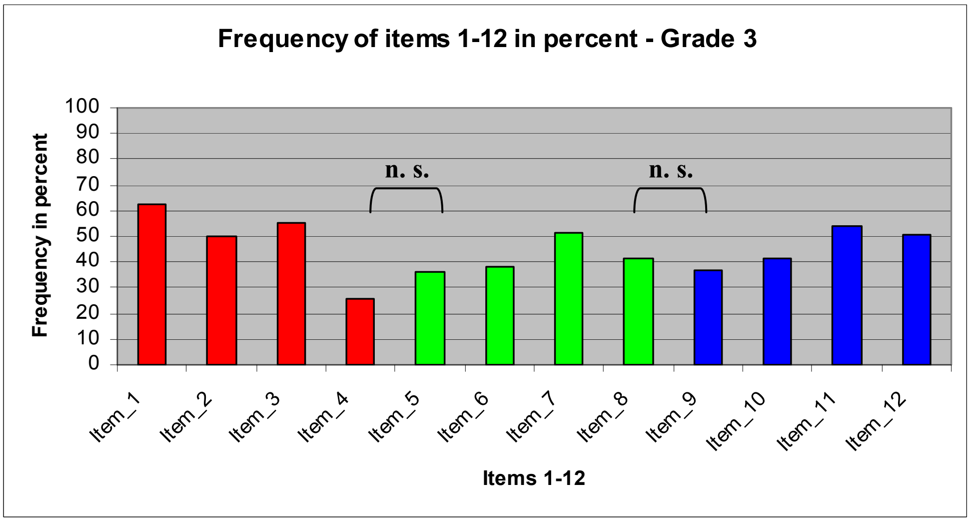


Figure 11 The frequency of items 1-12 in percent for the main task for the 3rd class

In grade 4, yet a different pattern occurred. Namely, in the second category there was build-up of PI and at the end the RPI showed up. Other than these, no significant change was found for both the PI and the RPI. The capacity of the memory did not increase significantly.

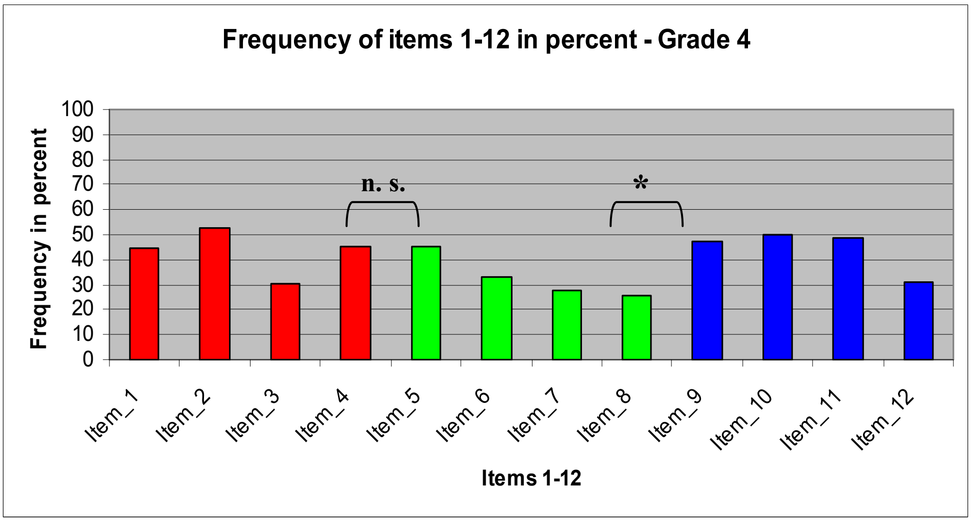


Figure 12 The frequency of items 1-12 in percent for the main task for the 4th class

Lastly, in grade 5, the initial pattern produced by the first graders was partially present again. Other than the first item of the last category, nearly perfectly all build-up of proactive interferences and release from proactive interferences were present. Consequently, it might be assumed that there were some processes between the 1st graders and the 5th graders that caused that many different patterns for the middle graders – the 2nd, the 3rd, and the 4th graders. Besides, the memory capacity of the 5th graders developed significantly.

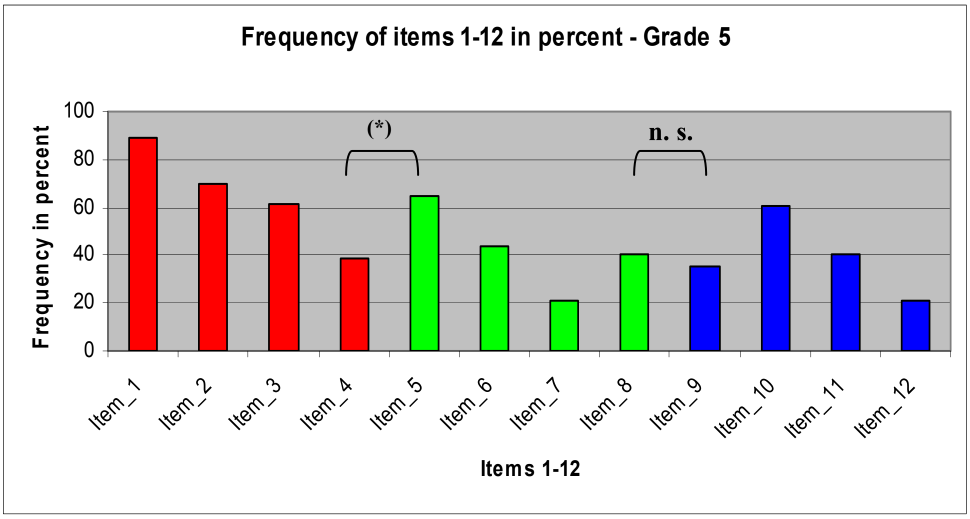


Figure 13 The frequency of items 1-12 in percent for the main task for the 5th class

The following figures (Figure 14-18) represent the development of the recall of the total number of items in relative order for each class (cx_y means the y^th item of the x^th category, i.e. c1_3 means the third item of the first category). The bars denote the absolute number of children who remembered the respective items in correct relative order. To calculate the relative order of items for one class, the order of each separate item in each category is summed up according to the answer of the subject. Note that the first suite is about the very general working memory capacity as it develops over grades (absolute number of all remembered items in a certain position). However, the following suite is about the amount of remembered items in correct relative order of items. This development shows more clearly how (relative) seriality develops.

The results for the relative order of items were very favorable for the first graders. The items remembered for each category approximately remained the same, i.e. the categories were recalled equally well. Also, both the PI and the RPI pattern showed up clearly for this grade. That means the first graders have already the categorization ability and present the full pattern for RPI. However, the overall performance of the class was low.

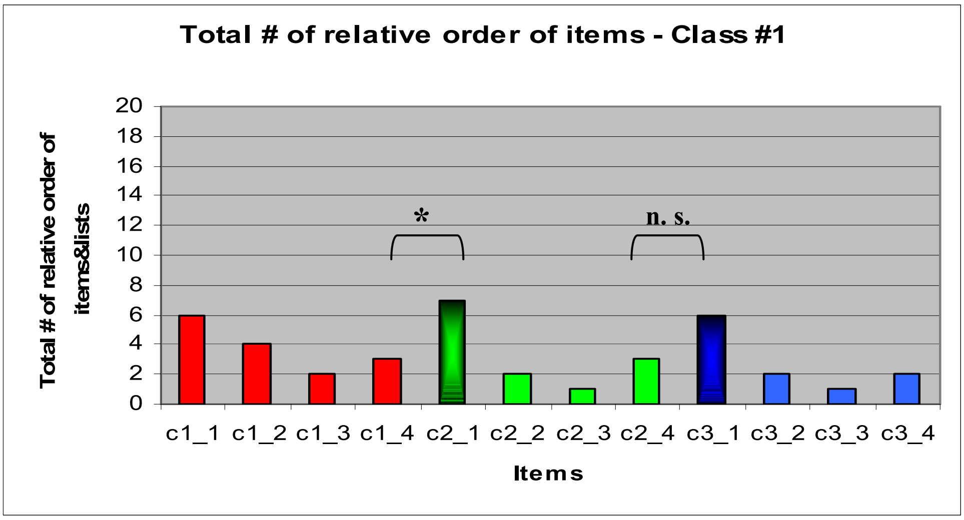


Figure 14 The relative order of items for the main task for the 1st class

The 2nd graders also remembered nearly the same number of items for each category. This time, the absolute level of the memory was higher. They presented the RPI pattern, as well. However, there seemed to have emerged a recency effect within the last category.

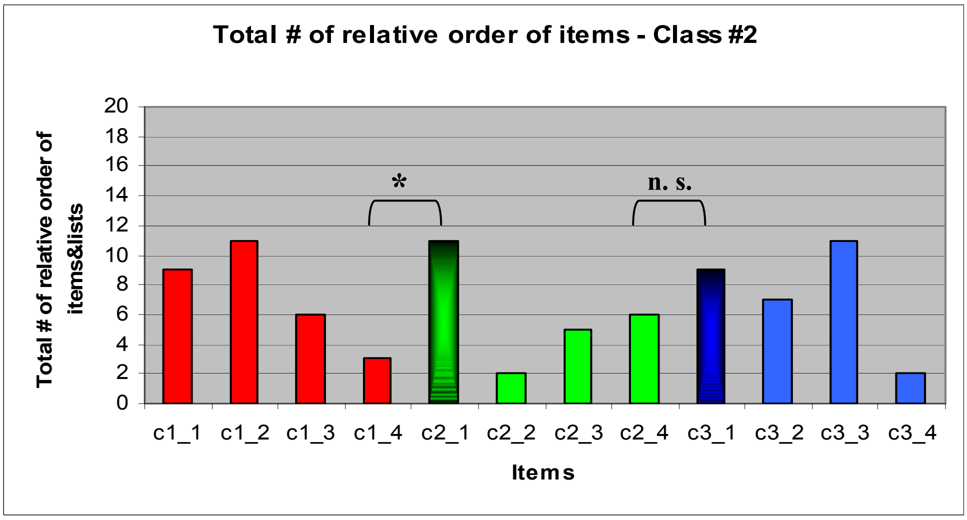


Figure 15 The relative order of items for the main task for the 2nd class

The performance of the 3rd graders was very bad, indeed. Possibly due to a memory reorganization, the performance of the children dropped. Although there was a primacy effect for the first category and the first item, the curves for the two other categories were quite flat. That is, the overall list was asymmetric. There was only one build-up of PI and subsequent RPI for the first category shift.

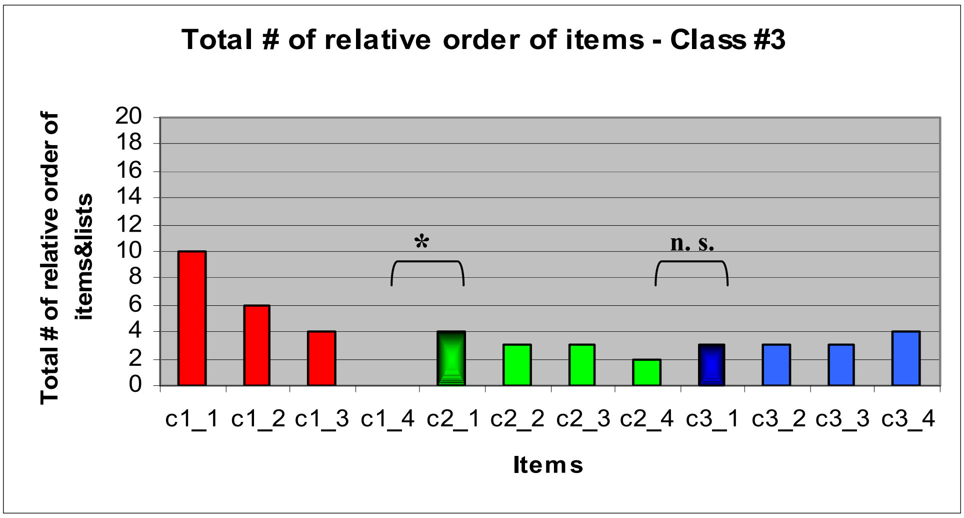


Figure 16 The relative order of items for the main task for the 3rd class

In grade 4, both the primacy effect got stronger for the first category and the general performance for remembering the items got better. However, like the 3rd graders, the categorical effects (PI and RPI) could not be seen in this figure. So the recall curve both for the second and the last category was quite flat as compared to the first category. However, as compared to grade 3, the overall memory capacity has increased. The RPI effect was present only for the second category shift. Again, the overall list was asymmetric and there emerged a primacy effect for the first category.

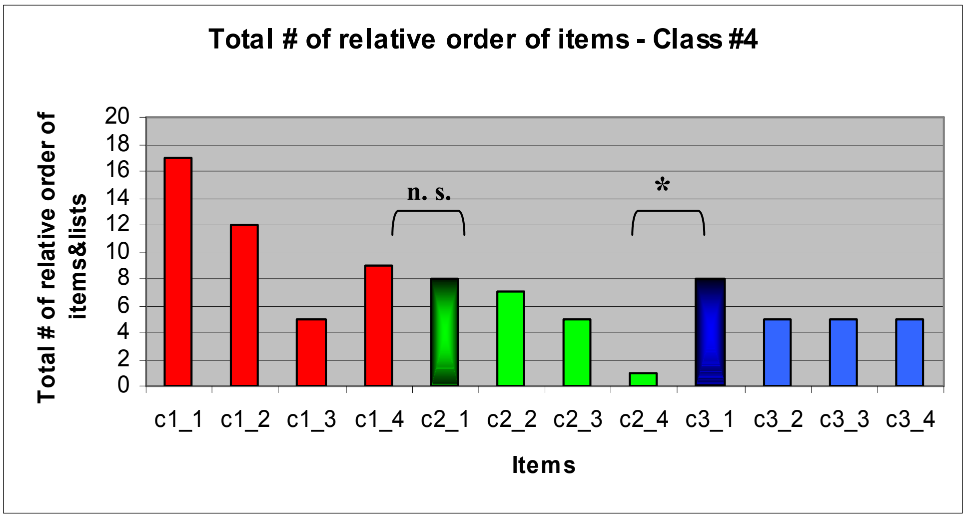


Figure 17 The relative order of items for the main task for the 4th class

In the 5th grade, the students' memory performance stabilized. Secondly, the RPI pattern that was reflected both in the 1st and the 2nd grade could not be observed for the children in 4th and 5th classes. These children's recall patterns mostly were symmetrical. However, in the 5th grade, there was both symmetry and asymmetry. The symmetry could be seen both at the beginning of the list and at the end of the list since the first and the last words were remembered much better than the middle words. As for the asymmetry, the primacy effect was stronger than the recency effect. For this grade, the primacy effect was just for the items; they could be observed in the first 1-2 items. However, the recency effect was for the last category. Lastly, the most interesting thing would be the "dip" in the middle probably indicating that the older children aimed to remember all items in list in the first place. Therefore, they did not consider the category shifts as much as the younger ones like the 1st and the 2nd graders. Then the list became asymmetric.

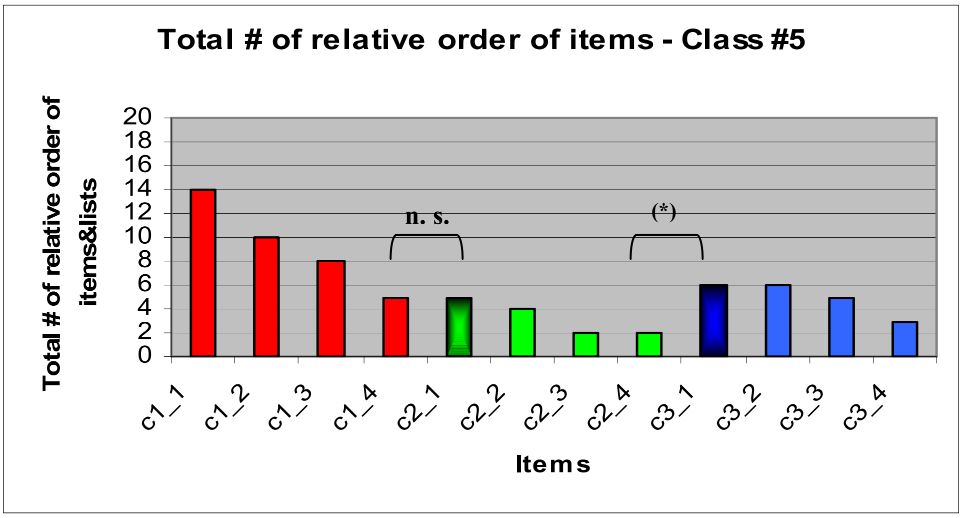


Figure 18 The relative order of items for the main task for the 5th class

To conclude, the overall results for the RPI effect show no development of the categorization ability. An asymmetry in overall list recall emerged, due to a primacy and recency effect for the first and last category/items. PI and RPI were present as early as grade 1, indicating that 6-yr old children can already categorize efficiently and use these abilities to organize their working memory processes. Besides, the stimulus list may not have been long enough for the 4th and 5th graders to create an RPI effect. Therefore, future studies that aim to study the RPI effect should use a 15 – word list instead of a 12 – word list.

4.1.2.1 Analyzing the RPI with ANOVA

We also ran an ANOVA in order to confirm the results of the non-parametric tests. We used repeated measures of ANOVA by selecting the first shift and the second shift within subject variables, separately. The between subject variables were grade and gender.

The first shift was significant in the overall sample ($F(1, 101) = 4.171$, $p = .044$). The first shift developed marginally significantly by grade class ($F(2, 275) = 4.171$, $p = .067$). The gender, and the interactions first shift * class, first shift *gender, and first shift * gender * class were all insignificant.

The second category shift did develop by the grade (F (1, 101) = 2.824, $p = .048$) but did not change due to gender. Other than the interaction second shift * gender * class (F (4, 101) = 3.009, $p = .022$), the remaining interactions were insignificant, i.e. second shift * class and second shift * gender.

All these results were also supported by the non-parametric tests' results.

4.1.2.2 Release from proactive interference and phonological similarity

In a short-term serial recall experiment, it was found that it is easier to remember the non-similar sounding items than the similar ones (Lewandowsky & Farrell, 2008). Therefore, phonological similarity within and across categories might be a candidate for explaining the insignificant results in the release from proactive interference. Such phonological similarity is inherent in item pairs such as inek-çilek, erik-inek, karpuz-kazak, kazak-kocuk, etc. Therefore, we established some criteria in order to compute this effect. An example list for calculating the phonological similarity score is shown below:

Example list produced by one child for calculating phonological similarity:

erik	(plum)
inek	(cow)
karpuz	(watermelon)
kazak	(sweater)
tavuk	(chicken)

We parsed the words in order to examine the number of identical phonemes with respect to their positions in the words. In our experiment, all words were 2-syllabic. A syllable has a maximum of three positions, i.e. the nucleus, the coda and the onset. The nucleus is the center of the syllable, which is the vowel. The coda comes after the nucleus and generally is a consonant. Lastly, the onset comes before the nucleus and usually is a consonant. We analyzed the similarities between these segments of the words in the produced lists.

Table 7 An example for analyzing the phonemes

1st syllable			2nd syllable		
Onset, C	Nucleus, V	Coda, C	Onset, C	Nucleus, V	Coda, Ca
k	a	r	p	u	z
k	a		z	a	k

While analyzing, only neighboring words were compared (i.e., erik-inek; …). In the below table, there is a total of 5 identical phonemes among the 4 consecutive comparisons (k-k, a-a, a-a, k-k, k-k). Therefore, the ratio is 5/4, i.e. 1.25. We expect more phonological similarity for the older children and less for the smaller ones. We base our prediction on findings from early language production in children. Thus, it has been shown in the developmental literature on slips of the tongue that phonological similarity becomes an organizing principle later in development as compared to semantic similarity, that is, only older children substitute words according to their phonological similarity (Jaeger, 2005).

Table 8 An example list for analyzing the phonological similarity effect

	e		r	i	k
	i		n	e	k
k	a	r	p	u	z
k	a		z	a	k
t	a		v	u	k

We also calculated the phonological similarity score for the empirical lists, that is, the lists that were presented to the subjects. The phonological similarity of the empirical list (Mean= .577, S.D. = .080) was significantly different from the phonological similarity (Mean= .771, S.D. = .362) of the list they produced themselves (Wilcoxon Signed Ranks Test (Z = - 4.657, $p < .001$). Overall, empirical phonological similarity was lower than the phonological similarity of the produced lists.

From this result, it can be inferred that subjects tend to remember the items more in terms of phonological similarity. They do not always remember the items in serial order but continue the serial order in terms of phonological similarity. That is, children tend to use phonological similarity as a memory strategy. However, the overall effect of the phonological similarity does not significantly change with age although the phonological similarity scores were somewhat higher for the older children than for the younger ones. In short, there was no change in phonological similarity over time (see Table 9). That means that any development in the sensitivity to phonological similarity must already have taken place prior to schooling (in line with Jaeger's findings). All children in our sample, irrespective of age, can make use of phonological similarity as a memory strategy.

Table 9 Descriptive statistics for the phonological similarity of items remembered by the subjects - gradewise

Grade	Mean	S.D.	N
1	,729	,423	20
2	,759	,365	24
3	,693	,255	16
4	,808	,305	22
5	,855	,435	19
Total	,771	,362	101

Furthermore, according to the results of a study on phonological similarity (Nimmo & Roodenrys, 2004), it was not beneficial when the first consonant and vowel were the same in the stimuli or the first and the last consonant were the same. Thus, the phonological similarity effect might not have affected the results at all. However, the difference between the empirical and the memory list shows that the subjects are sensitive to phonological similarity.

4.1.2.3 Release from proactive interference and gender

As for gender, there was no significant change for the release from proactive interference according to the Kruskal-Wallis Test results. More specifically, among male subjects, c1_4, c2_1, and c2_4 were not significant[2]. However, just c3_1 was significant (p = .032), but it was not very important since overall there was no significant effect of gender when it was used as between subject factor. Among female subjects, c1_4, c2_1, c2_4, and c3_1 were insignificant. Also, the results for the gender as between subject factor were insignificant.

4.1.3 Categorical cohesion

We also analyzed the categorical cohesion in the results of the subjects for the CFR-Test, that is, how many category (semantic) shifts there were in the remembered lists. In order to assess this measure the responses were examined according to the criterion "whether each following word belongs to the same or a different category". An example list for calculating the categorical cohesion score is shown below:

Example list from one subject for the calculation of the categorical cohesion:

	çorap	(sock)
	kazak	(sweater)
first category shift →	*kaplan*	*(tiger) False Memory!*
	inek	(cow)
	tavuk	(chicken)
second category shift →	çilek	(strawberry)

For this list, we measured two values, one with false memory items and one for without. To begin with the first measure (false memory (FM) items included), "çorap" and "kazak" belong to the same category, "clothes". Then, a new category, "animals", starts with "kaplan" (false memory) and continues with "inek" and "tavuk" from the same category. So there is a shift from one category to the other category. Next, "çilek" follows which is again from another category. This is the

[2] c1_4: the last item of the first category; c2_1: the first item of the second category; c2_4: the last item of the second category; c3_1: the first item of the last category

second category shift. In total we have 6 words. We use the following equation for the ratio or category shifts:

Ratio (+FM) = total number of category shift / total number of remembered items (+FM)

Therefore, for this example, the “Ratio+FM” is 0.33 (2/6) and “Ratio” is 0.4 (2/5).

We expect that the ratio should be higher for the smaller children since they should make more category shift than the older ones, that is, their output might not be as much organized by the categories than the output of the older children.

Since the data were not normally distributed (Kolmogorov-Smirnov Test and Shapiro-Wilk Test, $p < .05$), we used the Kruskal-Wallis Test. Neither the results for the Ratio+FM (Mean= .432, S.D. = .163) were not significant ($\chi^2(4) = 3.77$, $p = .437$) nor those without the false memories (Mean= .448, S. D. = .177; $\chi^2(4) = 2.19$, $p = .701$). The most sensible reason for the insignificant results might be the fact that the youngest subjects in the overall sample (aged 6 years) can use semantic categories as organizing principles of memory recall. That is, this capacity does not develop anymore in older children.

4.1.4 False memories

False Memories are erroneously remembered items that are related to the stimulus material semantically, e.g., “kaplan” (‘tiger’) in the “animals” category. A false memory score was calculated by summing up the incorrect words that the subjects uttered after having listened to the items (see Table 10). Since the normality test was significant (Kolmogorov-Smirnov Test and Shapiro-Wilk Test, $p < .05$), we used non-parametric tests. False memory decreases marginally over age according to the Kruskal-Wallis Test ($\chi^2(4) = 9.04$, $p = .06$). However, only between grade 1& 4 and 2&4, there was a significant difference ($p = .030$ and $p = .019$).

The below table represent the mean ranks values for the false memories for the whole sample. The marginal decrease appears between the 3rd and the 4th grade in which children produced less false memories.

Table 10 Absolute numbers of false memories across grades

	Grade	N	Mean Rank
Category Test – false memory	1	20	59,13
	2	24	59,19
	3	16	52,97
	4	22	39,98
	5	19	43,21
	Total	101	

Figure 19 demonstrates the differences between boys and girls according to the produced false memories. Seemingly, the boys always produced more false memories than the girls but not significantly. And the greater drop was between the 3rd and the 4th grade for both boys and girls alike.

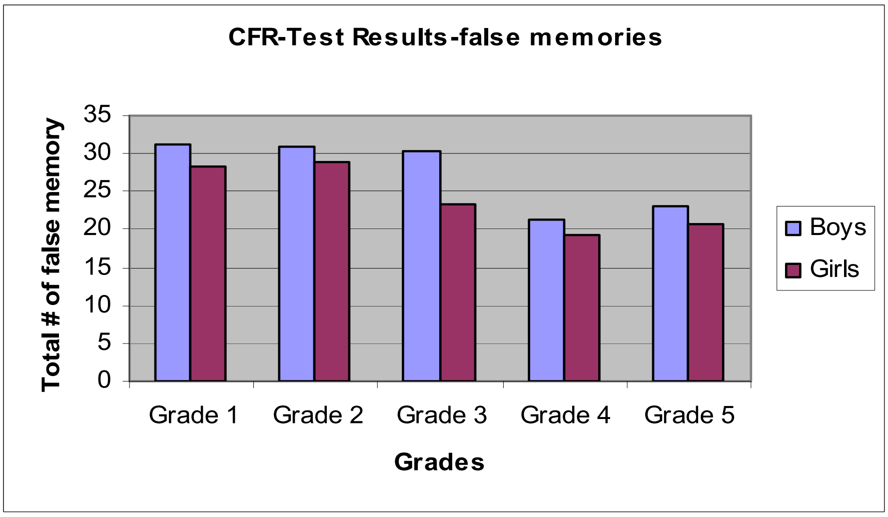


Figure 19 The development of false memories in the CFR-Test for girls and boys separately

It seems that the older children tend to have more veridical memories, while the younger ones are more prone to false memories. For example, in the CFR-Test, one of the 1st graders produced false memories such as “bilgisayar” (computer), “cep telefonu” (cell phone). These items were not in the list but in the physical environment of the subject. In support of this observation, Brainerd et al. (2002) asserted that "A core finding is that young children are often especially vulnerable to false memories, perhaps because verbatim memory for the actual events of their lives is poor." (p. 1363).

4.1.5 Analyzing the CFR-Test task with ANOVA

Non-parametric tests do not allow including more than one variable in each analysis, so that ANOVA had be used in order to explore any interactions between gender, category, items, and grade (for the mean table see Table 69 in Appendix D). Since we did all the other tests with non-parametric variants of ANOVA, we can compare the results between the parametric and non-parametric tests. Therefore, we analyzed the interactions between, grade, category, and items. Gender was analyzed in a separate ANOVA in order to keep the number of variables at a manageable number.

4.1.5.1 Category, grade and item interactions

The interactions between categories (3), items (4 for each category), and grade (5) were analyzed with ANOVAs. Grade was used as a between subject factor and category and items were within-subject factors in the mixed model. It was found that category and items were significant factors for the main CFR-Test ($F(2, 101) = 6.686$, $p < .001$; $F(3, 101) = 10.029$, $p < .001$). Also, the grade was found to be a significant factor ($F(4, 101) = 2.672$, $p = .037$) which indicated that overall the children's memory scores developed. The interaction between category and items was also significant ($F(6, 101) = 2.870$, $p = .009$). However, the other interactions were all insignificant (category*grade, $p = .171$; items*grade, $p = .441$; category*items*grade, $p = .081$).

Overall, the categories and the items had a main effect on the test. Namely, while the subject listened to items from the list, she would remember more items from the first category than from the second & the third categories and more items from the second category than the third one. This was also valid for the order of the items; the items at the beginning of each category were remembered better than the later ones. This reflects both the build-up of PI and the RPI effect.

4.1.5.2 Gender effect for the main task

In order to assess the effect of gender for the main task, we ran a second ANOVA. The dependent variable was the overall number of remembered items; grade and gender were used as between subject factors. The results were both insignificant for gender (F (1, 101) = 3.369, *p* = .070) and for the interaction between gender and class (F (4, 101) = 1.403, *p* = .239). Female subjects remembered only insignificantly more items (M=5.81, SD=1.14) than male subjects (M=5.34, SD=1.57). The results from the ANOVA confirm the earlier findings from the non-parametric tests. According to the results of the Kruskal-Wallis Test, among boys and girls, there was no significant absolute effect of gender, i.e. girls do not remember more items than boys overall (boys: $\chi^2(4) = 8.77$, $p = .067$; girls: $\chi^2(4) = 2.31$, $p = .679$) (see section 4.1.1 also for the non-parametric results on gender).

4.1.6 Discussion

According to the results of the CFR-Test, working memory developed in the children. That is, the ability to store information in memory improved during the primary school years. Other than the mere memory capacity as measured by the overall number of remembered items, the ability for serial recall as measured by the number of remembered items in absolute order, developed, as well. Serial recall has arguably an executive component since the items have to be lined up correctly in working memory. We found no significant gender effect between girls and boys which seems to suggest that there is no difference in the verbal abilities of the two genders, according to the results of the CFR-Test. However, this negative result is qualified by significant gender effects in development. If compared grade-wise, girls developed significantly in recall of the absolute order of items but boys (marginally) in the overall memory capacity. It might be the fact that at the beginning girls had already more improved memory capacities than boys so they did not show any further development for the overall memory capacity (see Figure 6). Again, for the boys, the ability to recall the absolute order of items did not develop further since it was relatively high already in the early grades.

The development of the RPI effect is very complicated, indeed. Overall, both category shifts were significant. However, this was not true for the different grades. The 1st graders did show a reliable pattern for the RPI effect. This means that the categorization ability is available already before school age. In the 2nd grade, only one significant category shift could be observed so the overall model started to change fast. Then, for the 3rd graders, a re-organization of memory might have occurred since the curve looked quite flat. This decrease in memory may indicate a change in the overall memory organization of the entire list. Also, in the 4th graders, no exact pattern of RPI could be seen. However, at the end, the 5th graders did show again the classical RPI pattern, at least in one category shift. Since the memory capacity for the older children also developed, their higher memory performance may have helped them overcome the interference effect so that no RPI effect could be seen due to their overall high performance. Thus, the ability to categorize and the organization of remembered items in terms of categories were already present in the 1st graders.

In order to explain these different RPI patterns, we investigated the phonological similarity and the categorical cohesion for the empirically observed lists of remembered items. However, neither of them showed any development over time and could therefore not explain the differences in the RPI changes. Thus, we stick to the conclusion that the RPI effect and the categorization ability are already present in the 1st graders and re-organization of memory might produce different kinds of RPI patterns across the older ages.

In addition to these, we found no developmental effect regarding the strength of the PI, since the interaction between the grade and the item was insignificant (see section 4.1.5.1). That is, the strength of the PI remains stable. Thus, this outcome is contrary to some findings in the literature, namely that either a decrease or an increase in PI occurs in development (Kail 2002). However, it is consistent with the results of other studies in the literature, namely that there is no change in PI in development (Kail (2002; see also section 2.3).

Moreover, we found that the youngest children had more false memories than the older ones, which is in line with the findings in the literature (cf. Brainerd et al., 2002). Thus the ability to resist likely but not actual information develops with grade.

4.2 Additional memory tasks

Gathercole (1999) found that generally, children increasingly well remember items over age, in almost all memory tasks. That is, the ability to solve these tests increases remarkably with age. So far, this finding is confirmed by the analysis of the CFR-Test (see Table 11 and Table 12). In addition, we also measured related memory abilities, with the WST, the WCST and the LST. The descriptive statistics of the three additional tasks are given in Table 12. In all three tests, there was a significant increase in performance over time (see Table 11).

Table 11 The test statistics of the Kruskal-Wallis Test for the additional tasks

	Word Span Test	**Wisconsin Card Sorting Test**	**Listening Span Test - level**	**Listening Span Test – absolute # of correct items**
Chi-Square	24,542	34,759	74,746	69,971
df	4	4	4	4
Asymp. Sig.	,000	,000	,000	,000

In the following, the results of all three additional memory tests will be presented in detail. Since the data was non-normally distributed for all tasks (K-S Test and Shapiro-Wilk Test was significant, $p < .05$), non-parametric tests were used.

Table 12 Descriptive statistics for the additional tasks (For the LST, two scores were calculated: (1) the attained set level of remembered items and (2) the absolute number of remembered items)

	Mean	S.D.	Median
Word Span Test	3,916	,784	4,00
Wisconsin Card Sorting Test	26,41	8,244	27,00
Listening Span Test – set level	2,045	,863	2,00
Listening Span Test – absolute # of correct items	7,594	4,771	6,00

4.2.1 The Word Span Test (WST)

The WST was clearly affected by grade (Kruskal-Wallis Test, $\chi^2(4) = 24.54$, $p < .001$, cf. Table 11). As shown in the below figure, the scores of the children increase linearly with their age, as the ones in Figure 20 by Gathercole (1999). The mean values and standard deviations for the WST for all grades are shown below.

Table 13 Descriptive statistics for the WST for all grades

	N	Mean	S.D.	Median
Grade 1	20	3.325	1.030	3,00
Grade 2	24	3.708	.690	4,00
Grade 3	16	4.125	.619	4,00
Grade 4	22	4.136	.468	4,00
Grade 5	19	4.368	.597	4,00

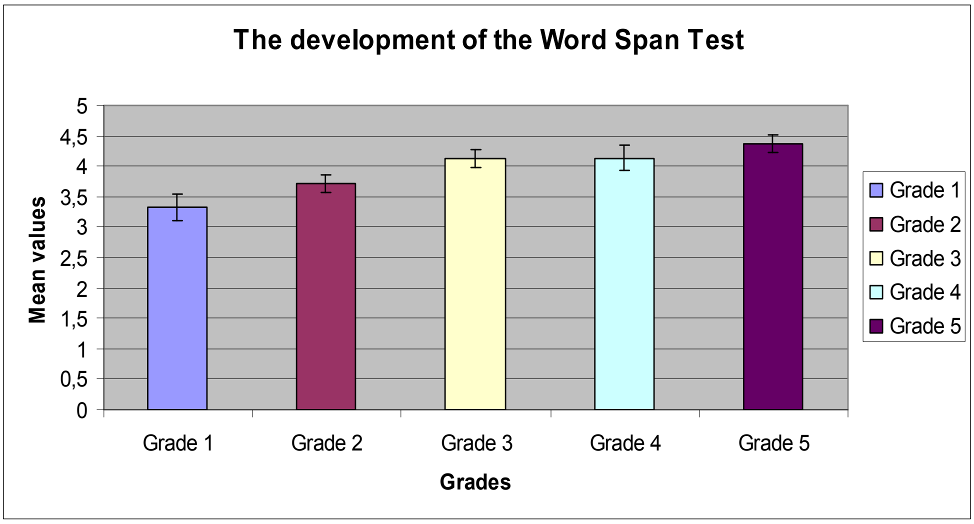


Figure 20 The development of the WST across grades (Error bars represents SEs)

In Figure 21, the development of the boys and girls can be observed regarding the WST. The developmental pattern for this task was very clear indeed. While girls develop steadily during this age range, the boys' scores jump both between the 1st and the 2nd grade and the 2nd and the 3rd grade. Then, they do not develop further.

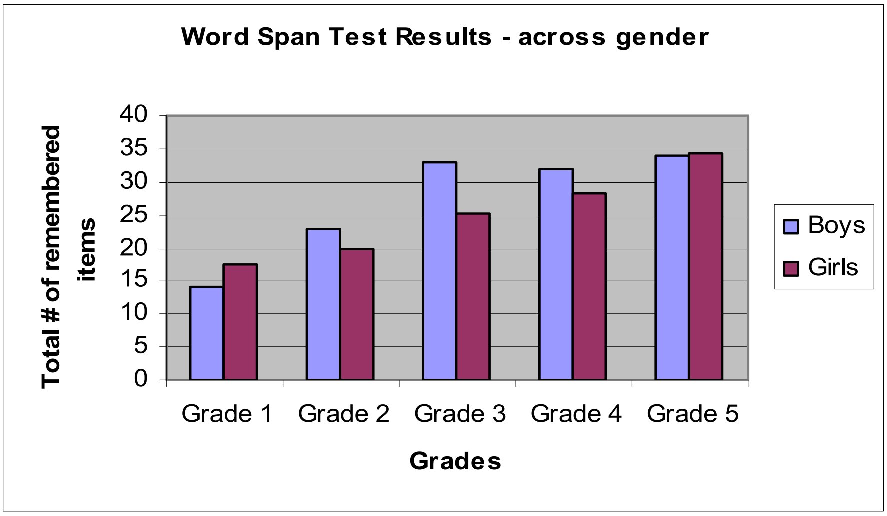


Figure 21 The development of the WST for girls and boys separately across grades

4.2.2 The Wisconsin Card Sorting Test (WCST)

There was also an age effect in the results of the WCST (Kruskal-Wallis Test, $\chi2(4) = 38.14$, $p < .001$, cf. Table 11). The table below represents the mean values and standard deviations for this task. In addition, Figure 22 shows that there was a linear increase for the WCST.

Table 14 Descriptive statistics for the WCST for all grades

	N	Mean	S.D.	Median
Grade 1	20	18.250	9.722	21,00
Grade 2	24	24.790	5.579	26,00
Grade 3	16	27.190	6.036	26,00
Grade 4	22	30.230	7.615	30,00
Grade 5	19	31.950	3.440	32,00

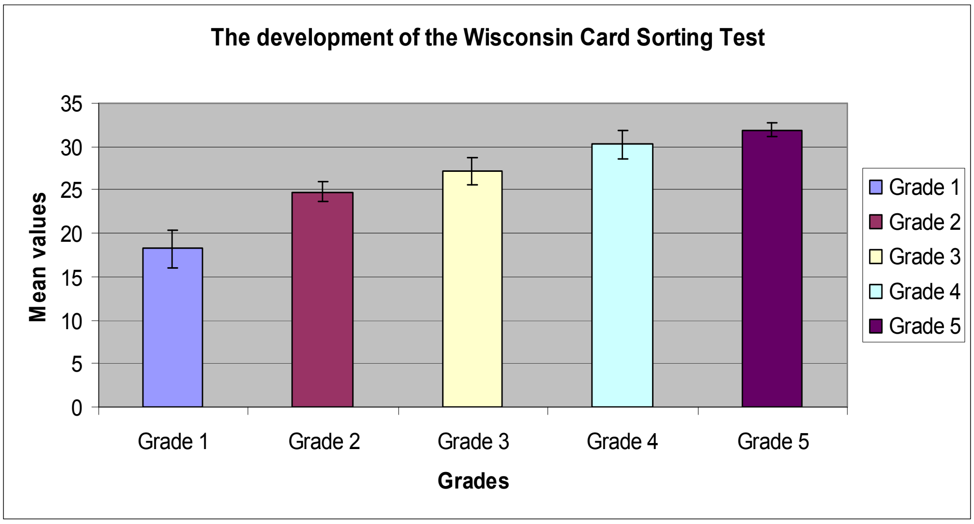


Figure 22 The development of the WCST across grades (Error bars represents SEs)

In the WCST, during the 1st and the 2nd classes, both boys and girls got nearly the same scores (see Figure 23). However, then boys' scores exceeded the girls' in the 3rd grade. Namely, the girls developed between the 1st and the 2nd grade and the 4th and the 5th grade. For the boys, the jumps were between the 1st and the 2nd classes and also between the 2nd and the 3rd classes. In the middle classes, boys were better than the girls but at the end girls' scores turned out to be higher.

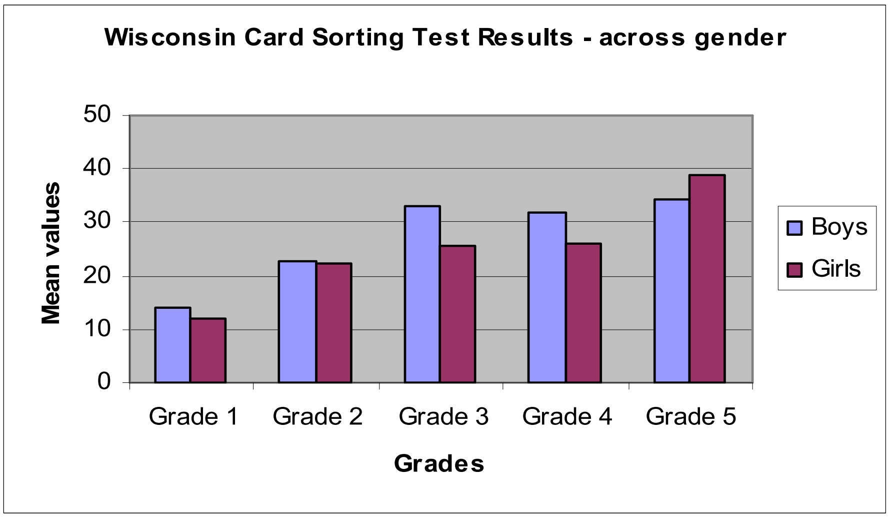


Figure 23 The development of the WCST for girls and boys separately across grades

4.2.3 The Listening Span Test (LST)

The LST was also affected by grade (Kruskal-Wallis Test, $\chi^2(4) = 74.71$, $p < .001$, cf. Table 11). However, this time, unlike the WST and the WCST, the development was step-wise and not linear.

Table 15 Descriptive statistics for the LST - level for all grades

	N	Mean	S.D.	Median
Grade 1	20	.700	.657	1,00
Grade 2	24	2.021	.312	2,00
Grade 3	16	2.000	.516	2,00
Grade 4	22	2.727	.369	2,75
Grade 5	19	2.737	.304	2,50

Table 16 Descriptive statistics for the LST – total # of recalled items for all grades

	N	Mean	S.D.	Median
Grade 1	20	2.050	1.877	3.00
Grade 2	24	5.542	2.146	5.00
Grade 3	16	6.188	3.209	5.00
Grade 4	22	12.045	3.709	11.50
Grade 5	19	12.053	2.345	11.00

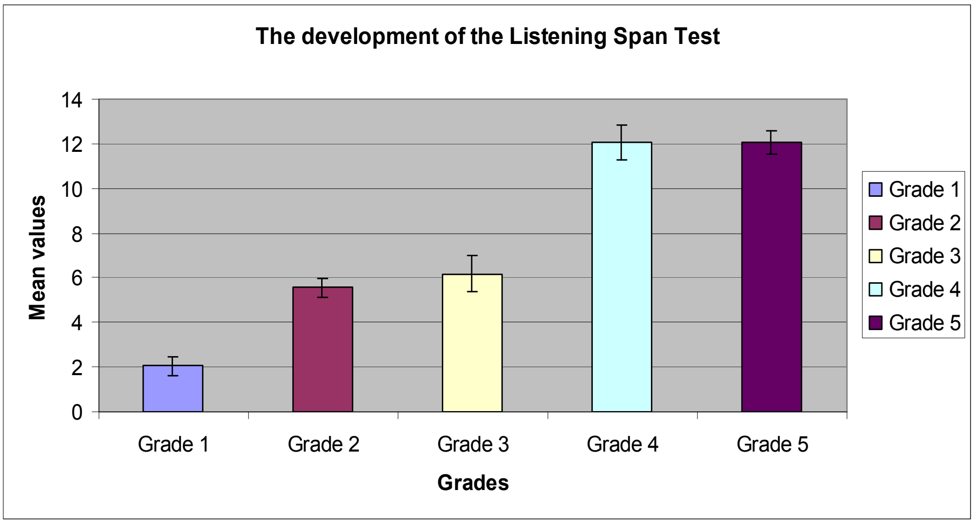


Figure 24 The development of the LST across grades (Error bars represents SEs)

Figure 25 reflects the nearly perfect developmental pattern of the LST. Namely, the first developmental increase was in between the 1st and the 2nd grade. Then the scores stabilized. However, between the 3rd and the 4th class, again there emerged a jump indicating a significant developmental increase in the task's results. Then, the results became stable again for the 5th class.

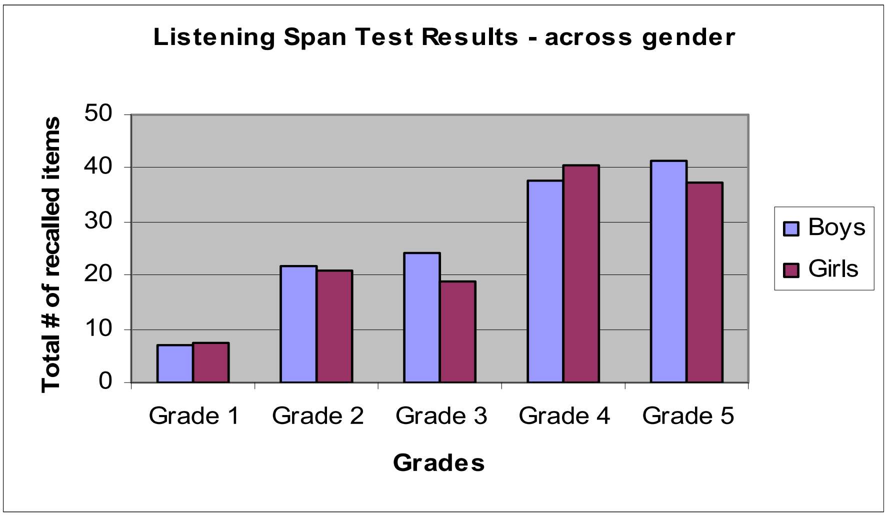


Figure 25 The development of the LST for girls and boys separately across grades

The following figure represents all additional tasks' results across the grades. It seems that both the WST and the WCST results develop in small amounts. However, in the LST case, the developmental increase is not kind of flat but step-wise.

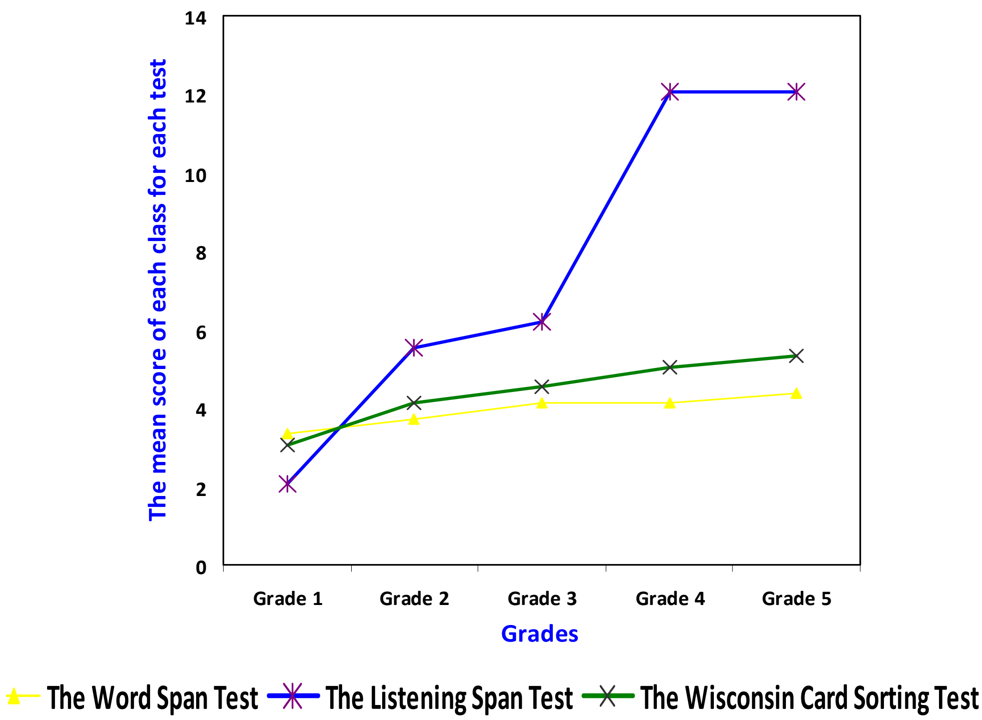


Figure 26 The development of all additional tasks across grades

4.2.4 Correlations between the tasks

We also looked at the correlations between the tasks. As a parametric correlation measure the Peterson's r and as a non-parametric correlation measure the Spearman's rho were used. The results of the parametric and non-parametric correlations were always convergent. Overall, there were high correlations between the tests that all reached significance (see Table 17). In the following, correlations for the various grades are reported. Specifically, for grade 1 and grade 5, there was only one significant correlation between the WST and the WCST (Pearson's r = .557, Spearman's rho = .483; and Pearson's r = .470, Spearman's rho = .507). For the second and third grades, the correlation between the WST and the LST was significant (Pearson's r=.463, Spearman's rho = .202 (p = .343, insignificant); Pearson's r=.685, Spearman's rho = .439 (p = .089, insignificant)). In grade 4, there were no significant correlations. Overall, the many significant correlations between the tasks were expected indeed since they all measure various (partly overlapping) aspects of working memory.

Table 17 Pearson's and Spearman's rho correlation table for the additional tasks for all subjects

	Pearson's r		**Spearman's rho**	
	WST	**WCST**	**WST**	**WCST**
WCST	.465**		.389**	
LST	.504**	.548**	.539**	.568**

Moreover, the presented results show that different working memory functions may develop at different times for children. For example, for the 1st graders, the relationship between the WST and the WCST indicates that both these two functions of the working memory started to develop at this age group or before this age. Then, the relation between the WST and the LST started to emerge for the 2nd and the 3rd graders. For this age group, the development of the functions for the WCST became stable but the functions related with the LST began to evolve. For the 4th graders, none of the relations was significant which indicates that there was nothing developing at this age group. Lastly, the relation between the WST and the WCST emerged again. Therefore, for these children, both the development of the LST and the development of the WCST re-appeared.

4.2.5 Gender effect for the additional tasks

We used the same statistical models, namely ANOVAs, as in section 4.1.5 to assess the effect of gender and age on the additional tasks (see Table 18 for the descriptive statistics). We obtained no significant effect for gender for any grades (the WST, $F(1, 101) = 3.033$, $p = .085$; the WCST, $F(1, 101) = .415$, $p = .521$; the LST, $F(1, 101) = 1.093$, $p = .299$). In addition, the class*gender interaction was neither significant for any grades (the WST, $F(1, 101) = .671$, $p = .614$; the WCST, $F(1, 101) = 1.518$, $p = .203$; the LST, $F(1, 101) = 1.925$, $p = .113$).

Table 18 The results of the additional tasks for the two genders: male and female

	Gender of the student	**Mean**	**S.E.**
Word Span Test	male	3,814	,098
	female	4,061	,102
Wisconsin Card Sorting Test	male	25,957	,946
	female	26,838	,989
Listening Span Test - level	male	1,984	,061
	female	2,102	,064
Listening Span Test – absolute # of correct items	male	7,349	,372
	female	7,912	,389

When we compared the results from the ANOVAs with their non-parametric counterparts, we could also observe that there was no change in the significance level between boys and girls for the additional tasks, i.e., the WST, the WCST, and the LST. Table 19 and Table 20 show for male and female subjects separately that there was a significant development in these tasks across grades. That is, boys and girls did not differ in their development in these tasks.

Table 19 The test statistics of the Kruskal-Wallis Test for the additional tasks for boys

	Word Span Test	Wisconsin Card Sorting Test	Listening Span Test - level	Listening Span Test – absolute # of correct items
Chi-Square	16,349	17,792	35,655	33,877
df	4	4	4	4
Asymp. Sig.	,003	,001	,000	,000

Table 20 The test statistics of the Kruskal-Wallis Test for the additional tasks for girls

	Word Span Test	Wisconsin Card Sorting Test	Listening Span Test - level	Listening Span Test – absolute # of correct items
Chi-Square	10,521	17,732	40,088	36,516
df	4	4	4	4
Asymp. Sig.	,033	,001	,000	,000

4.2.6 Discussion

The results revealed that children developed between the 1st and the 5th grade in the additional memory tasks. Firstly, for the WST, children's performance increased linearly in this task, in line with Gathercole (1999). Generally, the boys showed jumps between the first and the second classes and between the second and the third classes. However, the girls developed more gradually and slowly than the boys. Secondly, for the WCST, the increase was linear like for the WST. At the beginning, boys seemed to obtain better results than girls but then in the later classes the girls became better. Lastly, there was a step-wise development for the LST which indicated that this kind of ability did not increase yearly; instead it developed for a short period, then stabilized and then again developed (Gathercole, 1999). Although the developmental patterns look somewhat different for boys and girls, overall, no gender effect was found for any tasks. This means that for this age group, girls and boys develop similarly.

Furthermore, the correlations between the additional tasks were all significant. Specifically, in the 1st and the 5th grade, the correlation between the WST and the WCST was significant. This may hint at the underlying neuro-physiological and – anatomical development, namely that at the beginning and at the end of the primary school, the brain areas related with the WST and the WCST do develop but the LST is stable. For the 2nd and the 3rd graders, the relation between the WST and the LST is important. That is, during 8-9 years of age, the LST starts to develop along with an increase again in the WST. However, in the 4th grade, the development of all tasks has stabilized so that no correlation would show.

4.3 Multiple regressions

The results of the multiple regression test revealed that only the WST can predict the overall # of remembered items (the WST, t(100) = 3.434, p = .001, r = .414; the WCST, t(100) = 1.681, p = .096, r = .310; the LST, t(100) = -.857, p = .394, r = .194), when all predictors were entered at the same time.

The following table represents the correlations between the main task and the additional tasks. The WST and WCST are most dominant, but, the LST is also significantly related to the CFR-Test.

Table 21 The Pearson's correlation table for the CFR-Test and the additional tasks

	CFR-Test – total # of recalled items	
	Pearson's r	Sign.
Word Span Test	,414	,000
Wisconsin Card Sorting Test	,310	,001
Listening Span Test	,194	,026

In the multiple regressions, we used the WST, the WCST, and LST in one block. Our predicted variable was the CFR-Test (the total # of recalled items). In this model (see Table 22), the WST seems to explain all significant results for the main task. The predictive power of the WSCT falls short of significance, although it is also highly correlated with the CFR. However, the LST, although significantly correlated, fails to predict the main task. One possible reason could be that both the WST and the WCST explain one measure, i.e. the memory capacity or the executive functions. However, the LST comprises both working memory functions. Therefore, one big portion of the main task is explained by the WST and another smaller portion by the WCST. The remaining part to be explained is now very small. Thus nothing is left for the LST, indeed. Besides, the partial and the part correlations also indicate how the overall variance of the main tasks is to be explained by the other tasks. The partial and the part correlation for the WST did not drop significantly, while the WCST lost half of its correlation. For the LST, the situation was even more dramatic, i.e., it dropped to zero.

Table 22 The multiple regression results for the CFR-Test (total #of recalled items) and the additional tasks

Model						Correlations		
		B	SE B	ß	Sig.	Zero-order	Partial	Part
1	Constant	2,321	,691		,001			
	Word Span Test	,668	,194	,375	,001	,414	,329	,313
	Wisconsin Card Sorting Test	,032	,019	,190	,096	,310	,168	,153
	Listening Span Test - absolute # of correct items	-,029	,034	-,099	,394	,194	-,087	-,078

Note R Square Change = .195 for Step 1.

The below collinearity table represents some data about the three additional tasks in terms of their linear relatedness. In the table, the WST loads highly (92%) on a single dimension, while the WCST loads highly (99%) on a different dimension. This means that both the WST and the WCST can explain only one independent measure separately which are the memory capacity and the executive functions, respectively. However, the LST shares proportions with the other task and accumulates only 72% on its own dimension. This underlines the somewhat hybrid character of the LST as a measure of complex working memory.

Table 23 The collinearity diagnostics results for the CFR-Test and the additional tasks

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				Constant	Word Span Test	Wisconsin Card Sorting Test	Listening Span Test - absolute # of correct items
1	1	3,765	1,000	,00	,00	,00	,01
	2	,176	4,624	,04	,01	,00	,72
	3	,042	9,428	,10	,07	,99	,14
	4	,016	15,164	,86	,92	,00	,13

We also ran a stepwise multiple regressions in which we changed the order of tasks so that now the WST would only be entered last. This time, there were three blocks in the following order:

1. WCST
2. LST
3. WST

For the first block, we found that the WCST could explain the main task significantly. However, when the second task, i.e., the LST, was entered in the second block, the WCST lost some of its power, while the LST remained insignificant. Lastly, in the third step, when the WST was introduced, it absorbed all variance proportions possessed by the WCST and the LST before.

Table 24 The multiple regression results for the CFR-Test and the additional tasks (with a different blocking order)

Model		B	SE B	ß	Sig.
1	Constant	4,178	,447		,000
	Wisconsin Card Sorting Test	,053	,016	,310	,002
2	Constant	4,186	,450		,000
	Wisconsin Card Sorting Test	,049	,019	,291	,013
	LST - absolute # of correct items	,010	,034	,034	,765
3	Constant	2,321	,691		,001
	Wisconsin Card Sorting Test	,032	,019	,190	,096
	LST - absolute # of correct items	-,029	,034	-,099	,394
	Word Span Test	,668	,194	,375	,001

Note R Square Change = .096 for Step 1; R Square Change = .001 for Step 2; R Square Change = . 098 for Step 3.

Additionally, we changed the predicted value to the absolute order of items in the multiple regressions. We did this in order to find out whether different working memory factors would underlie serial working memory as compared to the overall working memory capacity. Table 25 shows the correlations between the predicted value and the other tasks. Now, the LST turned out to be the most powerful predictor of the main task. Then, it was followed by the WCST and the WST. The reason for this outcome might be that the absolute order of items factor includes an executive part. Because it does not just require keeping the information in the memory, it also demands manipulation of information. Since the LST includes both the storage of the information and the maintaining the information, the absolute order of items factor went along with the results of the LST.

Table 25 The Pearson's correlation table for the CFR-Test (the absolute order of items) and the additional tasks

	CFR-Test – absolute order of items	
	Pearson's r	Sign.
Word Span Test	,208	,019
Wisconsin Card Sorting Test	,262	,004
Listening Span Test	,319	,001

Also, the multiple regressions table supports the significance of the LST as a predictor of serial working memory. It was the LST which was (marginally) significant. For the partial and the part correlations, both the values for the WST and the WCST disappeared while for the LST it was still strong.

Table 26 The multiple regression results for the CFR-Test (the absolute order of items) and the additional tasks

Model						Correlations		
		B	SE B	ß	Sig.	Zero-order	Partial	Part
1	Constant	,026	,595		,966			
	Word Span Test	,049	,168	,034	,770	,208	,030	,028
	Wisconsin Card Sorting	,016	,016	,116	,329	,262	,099	,094
	Listening Span Test - absolute # of correct iter	,057	,029	,238	,053	,319	,195	,188

Note R Square Change = .113 for Step 1.

Lastly, more specifically, in grade 1, just the WST can predict the result of the Category Test (total # of recalled items). In the other grades, none of the tests can predict the main test (see Table 27).

Table 27 Multiple regression results for the additional tasks

	WST		*WCST*		*LST*	
	r	Sign.	r	Sign.	r	Sign.
Grade 1	.674	**.017***	.480	.520	.217	.948
Grade 2	.210	.188	-.263	.391	-.171	.286
Grade 3	-.309	.152	.104	.836	-.052	.422
Grade 4	-.018	.858	.062	865	.097	.682
Grade 5	.355	.076	-.048	.346	-.172	.430

4.3.1 Discussion

According to the results of the multiple regressions only the WST could significantly predict the main task, in terms of the total # of recalled items. Since both the WST and the CFR-Test relied on overall memory capacity, the influence of the WST on the CFR-Test was easily predictable. Other than this, not only the WCST but also the LST could predict the main task. However, since the underlying mechanisms of these tasks differ somewhat, the effects disappear when the other tasks come into play - only the WST still could predict the main task. The executive functions are measured by the WCST but the LST is not just a measure of the executive functions but also of the memory capacity. Therefore, since some common resources were shared by the different tasks, when all 3 experiments were entered into the multiple regressions, they could not explain the main task better as each single one could do in a simple linear regression. It turned out that the WST had the strongest predictive ability overall.

When the criterion was changed, namely, when the absolute order of items was to be predicted, all tasks were significantly correlated with the criterion, again. However, this time the LST became the most important one in the hierarchy. Since the ability to store items in mind serially needs manipulation of information in memory, it does require not only memory capacity but also some executive functions. Among the additional tasks, the LST shared variance with the WCST on executive functioning as well as with the WST on memory capacity. This underlines its hybrid character as a measure of complex working memory.

4.4 Teacher-related measures

We also asked some additional questions to the teachers of the subjects. We had a 5-point scale (very low - low – medium - good - very good) for teachers to score the related questions for the students (see Table 28 for the descriptive statistics). The questions are the following ones (see Appendix C for the questionnaire):

1. How intelligent do you consider the student?
2. How well can the student memorize?
3. How well can the student understand a given topic?
4. How efficiently can the student use his/her own language?

According to the results of Tillman et al. (2008), all four components – verbal- and visuospatial short-term storage and verbal- and visuospatial executive processes - of the working memory were found to be strongly and independently related with intelligence. Thus, it means that both the storage and the executive processes of the working memory are related to the intelligence of the children. Therefore, we could compare the test results with the teachers' ratings of the children's intelligence.

The judgments of the teachers on the children's performances could, of course, not be considered as the objective measures of their intelligence. Thus, it is possible to get insignificant correlation between the judgments and the memory task results since the teachers have only an intuitive access to intelligence.

Both descriptive statistics and the Pearson's and Spearman's correlation table are shown in Table 29. All the measures seem to be highly redundant. Therefore, we will only focus on the "intelligence" measure in the future.

Table 28 Descriptive statistics for the teachers' opinions

	Mean	**S.D.**
Teacher's opinion 1 – "intelligence"	3,46	,911
Teacher's opinion 2 – "memory"	3,43	,909
Teacher's opinion 3 – "understanding"	3,47	,923
Teacher's opinion 4 – "language"	3,34	,852

Table 29 Pearson's and Spearman's correlation table for the teachers' opinions

	Pearson's r			**Spearman's rho**		
	Mem.	Und.	Lang.	Mem.	Und.	Lang.
Intelligence	,862**	,840**	,754**	,855**	,857**	,749**
Memory		,789**	,756**		,804**	,745**
Understanding			,778**			,801**

We looked further for the correlations between the main task and the "intelligence" measure. Table 30-34 represents these correlations for intelligence*overall number of remembered items, intelligence* absolute order of items & lists, intelligence*all additional memory tasks (the WST, the WCST, the LST).

As can be seen from the below table that there was a zero or slightly negative but insignificant correlation between the overall number of remembered items and the teacher's evaluation of "intelligence" (Pearson's r = -.023, p = .817; Spearman's rho = -.114, p = .255).

Table 30 Pearson's and Spearman's correlation table for intelligence and overall number of remembered items

	CFR-Test – total #of recalled items			
	Pearson's r	Sign.	Spearman's rho	Sign.
Intelligence	-,023	,817	-,114	,255

The correlation between absolute order of items & lists and the intelligence was not significant either (Pearson's r = .158, *p* = .114; Spearman's rho = .157, *p* = .117).

Table 31 Pearson's and Spearman's correlation table for intelligence and absolute order of items & lists

	CFR-Test – absolute order of items&lists			
	Pearson's r	Sign.	Spearman's rho	Sign.
Intelligence	,158	,114	,157	,117

For the additional tasks, only the correlation between the WCST and the teachers' rating of intelligence was significant (Pearson's r = .212, *p* = .033; Spearmen's r = .199, *p* = .046). However, the correlation for the WST (Pearson's r = .180, *p* = .072; Spearman's rho = .156, *p* = .120) and the LST (Pearson's r = .179, *p* = .074; Spearman's rho = .174, *p* = .082) were not significant (see Table 32-34).

Table 32 Pearson's and Spearman's correlation table for intelligence and the WST

	Word Span Test			
	Pearson's r	Sign.	Spearman's rho	Sign.
Intelligence	,180	,072	,156	,120

Table 33 Pearson's and Spearman's correlation table for intelligence and the WCST

	Wisconsin Card Sorting Test			
	Pearson's r	Sign.	Spearman's rho	Sign.
Intelligence	,212	,033*	,199	,046*

Table 34 Pearson's and Spearman's correlation table for intelligence and the LST

	Listening Span Test			
	Pearson's r	Sign.	Spearman's rho	Sign.
Intelligence	,179	,074	,174	,082

Finally, we computed the correlation between the overall remembered items * the additional memory tasks and absolute order of items & lists * the additional memory tasks with and without the controlling for the variable "intelligence". For the interaction between the overall remembered items * the additional memory tasks, except the LST (Pearson's r = .194, p = .052; Spearmen's r = .221, p = .026), all additional tests were highly correlated with the overall remembered items (the WST, Pearson's r = .414, p = .000; Spearmen's r = .288, p = .004; the WCST, Pearson's r = .310, p = .002; Spearmen's r = .205, p = .040) (see Table 35).

Table 35 Pearson's and Spearman's correlation table for the overall remembered items and the additional memory tasks

	CFR-Test – total #of recalled items			
	Pearson's r	Sign.	Spearman's rho	Sign.
WST	,414**	,000	,288**	,004
WCST	,310**	,002	,205**	,040
LST	,194	,052	,221*	,026

For the interaction between the absolute order of items & lists * the additional memory tasks, except the WCST (Pearson's r = .211, p = .034; Spearmen's r = .177, p = .076), all additional tests were highly correlated with the absolute order of items

& lists (the WST, Pearson's r = .236, p = .018; Spearman's rho = .208, p = .037; the LST, Pearson's r = .214, p = .031; Spearman's rho = .261, p = .008) (see Table 36).

Table 36 Pearson's and Spearman's correlation table for the absolute order of items & lists and the additional memory tasks

	CFR-Test – absolute order of items&lists			
	Pearson's r	Sign.	Spearman's rho	Sign.
WST	,236**	,018	,208**	,037
WCST	,221**	,034	,177	,076
LST	,214**	,031	,261**	,008

For the partial correlations, we first controlled for the "intelligence" while correlating for the overall remembered items & lists * the additional memory tasks. As shown in Table 37, there was no significant change in the correlation values. The correlation between the LST and the main task, however, was significant this time (r = .201, p = .045). The other tasks were again significantly correlated with the main task (the WST, r = .425, p = .000; the WCST, r = .322, p = .001).

Secondly, we controlled for the "intelligence" while correlating for the absolute order of items & lists * the additional memory tasks. The partial correlation table, shown in Table 37, reflects similar results as in Table 35 and Table 36. The correlation between the WCST and the absolute order of items & lists was insignificant (r = .184, p = .067). Still, the other tasks were highly significantly correlated with the main task (the WST, r = .213, p = .033; the LST, r = .279, p = .005).

To conclude, we can say that controlling for "intelligence" does not change the correlations.

Table 37 Partial correlation table for the absolute order of items & lists * total #of recalled items and the additional memory tasks controlling for the intelligence

	Intelligence					
	WST		**WCST**		**LST**	
	Corr.	Sign.	Corr.	Sign.	Corr.	Sign.
CFR-Test – absolute order of items&lists	,213*	,033	,184	,067	,279**	,005
CFR-Test – total #of recalled items	,425**	,000	,322**	,001	,201*	,045
WST			,444**	,000	,487**	,000
WCST					,530**	,000

4.4.1 Discussion

Overall, the teacher-related measures were significantly correlated with each other. This might be due to the fact that the judgments of the teachers on these questions covered the general performance of each child in the class. Thus, separating the general performance of the child into different categories like intelligence, memory, etc. was not very helpful, indeed. Also, the subjectivity of the teachers might have an effect on these judgments.

Furthermore, the correlation between the intelligence measure and the total # of recalled items (also the absolute order of items) was insignificant. Thus, it seemed that these measures were not investigating the actual memory capacity in children. However, when the additional tasks came into play, then for the WCST a significant correlation was found. Therefore, what the teachers could judge in their students might be related to the simple executive functions (not the harder and more complex functions as measured in the LST). Since the children were 1^{st} – 5^{th} graders, their elementary executive memory capacity might have developed already while the more complex functions, which the teachers are not able to judge, are still developing.

Moreover, controlling for teacher's judgments of "intelligence" did not change the results for any task. Thus, this also indicated that the teacher-related measures did not tap the crucial components of memory such as working memory capacity; rather, they covered very basic opinions of the teachers about the cognitive ability of their students.

4.4 Relations between Socio-Economic Factors and the CFR-Test as well as the Additional Tasks

To assess the relation and possible influence of socio-economic factors on all tasks, data on the education level of the father, father's occupation, the marriage status of the parents, the number of siblings of the pupil, and the order of birth were collected for each subject (see Appendix C for the questionnaire). Education level and occupation of the mother did not vary within the overall sample. Mothers had predominantly primary school education and were housewives. Therefore, these variables were not taken into consideration. However, for fathers these two variables showed sufficient variation as to be used as a predictor in the subsequent correlation and multiple regression analyses.

The education level of the father was coded as follows:

1. none
2. primary school
3. secondary school
4. high school
5. university

The father's occupation was coded as follows:

1. no job
2. farmer
3. worker
4. worker-abroad
5. tradesman
6. civil servant
7. other

The marriage status of the parents was coded as follows.

1. married, living together
2. divorced
3. married, father working abroad

In an attempt to find out whether these socio-economical and family variables are related and possibly have an impact on the memory performance of the subjects, correlations and subsequently, multiple regressions were run with the former variables as predictors and the latter variables as dependent variables.

In multiple regressions, the set of independent variables were organized into three blocks, serving as regression models:

I. the education level of the father and father’s occupation
II. the marriage status of the parents
III. the number of siblings or the order of birth

Within the first model (block), the two predictors were entered simultaneously (“Enter” method). Within the third model (block), we either chose “number of siblings” or “order of birth” but never both together, in order to avoid any confounding). Since both the education level of father and father’s education are the main factors that influence the results of the tests, they were chosen as the first block members. For the next block, the marriage status has the secondary importance since there are several factors that interfere with the marriage status of the family. Thus, it cannot be easily analyzed. Lastly, both the number of siblings factor and the birth order factor are the sub-factors inside the general family factors.

The following tables represent the results of the correlations and multiple regressions between all tasks and the socio-economic and family factors separately.

In the multiple regression tables, B represents the regression coefficient (slope of the line) which predicts the change in dependent variable and S.E. B. shows the standard deviation value for the regression coefficient. The beta value reflects the strength of each predictor over the predicted value. The constant value for each variable indicates the Y-intercept of the regression function. R Square Change represents the strength between the predictor variable and the criterion variable.

Firstly, it turned out that the socio-economic and family variables were not related to the CFR-Test (see Table 38, all multiple regression results for the CFR-Test were insignificant and are therefore not reported). The main task just required the short-term memory of the participant, i.e., they should listen to the 12 words and remember them in order. Therefore, it might be inferred that no socio-economic factor play a role in such an experiment.

Table 38 Pearson's correlation table for the CFR-Test and socio-economic factors

	Categorical Free Recall Test	
	Pearson's r	Sign.
The education level of the father	-,026	,398
The father's occupation	-,001	,496
Married or divorced or abroad	-,071	,240
The number of siblings	,115	,125
The order of birth	,083	,204

However, the WST was related to two factors, namely, the education level of the father and father's occupation (see Table 39). Since all mother-related factors were constant, we may assume that the factors like the education level and the occupation of the father had some effects on the WST performance. Similarly, Bjerkedal (2007, p.510) also studied the influence of parental education level, however, not the father's but the mother's education level. He indicated that "*The mean of standardized scores for GA ("general ability" score indicating general intellectual score, addition by G.Ü.) of second born men was significantly lower than the mean of scores for first born for all levels of mother's education*." However, this study was just performed with young Norwegian men. Therefore, these results cannot be directly compared with ours.

One of the effects might be that during school years the increase of the children's vocabularies is fostered by the educational level of the father. Since the education level is directly related with one's vocabulary size, those children with fathers having a high education level may get better results than others. The other factor was the father's occupation which reflects the family income (since all mothers were housewives). The significant correlation between this factor and the WST indicated that the father's job was related to the vocabulary size of the children –probably because of different social environments due to the different occupations of fathers. In addition to these correlative relations, the result of multiple regressions supported the effect of the father's occupation.

Table 39 Pearson's correlation table for the WST and socio-economic factors

	Word Span Test	
	Pearson's r	Sign.
The education level of the father	,167	**,048****
The father's occupation	,230	**,010****
Married or divorced or abroad	,015	,440
The number of siblings	,122	,113
The order of birth	-,029	,387

Table 40 Multiple regression results for the WST and socio-economic factors

Model		B	SE B	ß
Step 2	Constant	3,317	,346	
	The father's occupation	,130	,065	0,221*
Step 3	Constant	3,013	,415	
	The father's occupation	,129	,065	0,220*

Note R Square Change = .003 for Step 2; R Square Change = .016 for Step 3, *p = .050.

The results of the WCST were affected by two other factors, namely the marriage status of the parents and the number of siblings of the participants (see Table 41). If the father of the child is out of city or country for a long time or the parents were divorced, then the performance of the child drops significantly. Therefore, it is beneficial to live together with both mother and father according to the negative correlation of marriage status. The number of siblings was positively correlated with the WCST, i.e., if one has more siblings then one performs better than others with fewer siblings. One possible reason might be that the number of brothers/sisters creates more communication between siblings so that experiments like the WCST may be positively affected by such relations. Likewise, the result of multiple regressions revealed that the marriage status of the parents and the number of siblings were important.

Table 41 Pearson's correlation table for the WCST and socio-economic factors

	Wisconsin Card Sorting Test	
	Pearson's r	Sign.
The education level of the father	,120	,116
The father's occupation	-,058	,282
Married or divorced or abroad	-,256	**,005****
The number of siblings	,295	**,001****
The order of birth	,102	,154

Table 42 Multiple regression results for the WCST and socio-economic factors

Model		B	SE B	ß
Step 2	Constant	26,853	3,617	
	Together or divorced or abroad	-2,841	1,209	-0,249*
Step 3	Constant	19,531	4,170	
	Together or divorced or abroad	-2,725	1,157	-0,239*
	The number of siblings	1,950	,619	0,294*

Note R Square Change = .053 for Step 2; R Square Change = .086 for Step 3, *p < .050.

Lastly, there were three factors regarding the LST results: the education level of the father, the number of siblings and the order of birth (see Table 43). The effect of the new factor – the order of birth– can be interpreted such that younger siblings should listen to the order of the older ones and older ones can give commands to the younger ones. If you are younger then you learn to execute the various commands of the older siblings. So the communications with you and your siblings would help you in such tasks. Similarly, the multiple regression results show the same results with the correlations (see Table 44 and 45).

Table 43 Pearson's correlation table for the LST and socio-economic factors

	Listening Span Test	
	Pearson's r	Sign.
The education level of the father	,293	**,001****
The father's occupation	,139	,082
Married or divorced or abroad	-,102	,156
The number of siblings	,314	**,001****
The order of birth	,184	**,033****

Table 44 Multiple regression results for the LST and socio-economic factors (number of siblings)

Model		B	SE B	ß
Step 1	Constant	1,902	1,901	
	The education level of the father	2,348	,857	0,276*
Step 2	Constant	2,883	2,063	
	The education level of the father	2,157	,869	0,253*
Step 3	Constant	-1,801	2,347	
	The education level of the father	2,302	,822	0,270*
	The number of siblings	1,247	,349	0,325*

Note R Square = .089 for Step 1; R Square Change = .014 for Step 2;
R Square Change = .106 for Step 3, *p < .050.

Table 45 Multiple regression results for the LST and socio-economic factors (order of birth)

Model		B	SE B	ß
Step 1	Constant	1,902	1,901	
	The education level of the father	2,348	,857	0,276*
Step 2	Constant	2,883	2,063	
	The education level of the father	2,157	,869	0,253*
Step 3	Constant	1,094	2,189	
	The education level of the father	2,299	,856	0,270*
	Birth order of the subject	,653	,303	0,204*

Note R Square = .089 for Step 1; R Square Change = .014 for Step 2;
R Square Change = .041 for Step 3, *p < .050.

4.4.1 Gender effect

We also looked for a gender effect regarding the socio-economic effects. Firstly, there was no significant difference between girls and boys according to the CFR-Test (see Table 49 and 50), as evidenced by the insignificant correlations and the multiple regressions. However, for the boys, the number of siblings affected their performance. Importantly, there were negative relation for girls and positive relation for boys. Thus, taken together, there arose big difference between girls and boys in this respect. One possible reason for this might be that in small villages like in Yozgat, girls are somewhat limited in their mobility as they grow up, and this may influence their development. Girls are frequently forced to stay with their mothers at home, do housework, etc. However, boys can free to roam around the village, play till midnight. Also, if a boy has several siblings, then he can take advantage of their environment, friends, physical strength, etc. in order to become stronger among his own friends. However, girls cannot take the advantage of having several siblings since her freedom is more restricted. Thus, the number of siblings factor may have a relation to "executive function" that plays role when a child's mobility is limited or not. In other words, if a child is not allowed to express and act according to his/her own wishes, his/her executive functions might become repressed.

Table 46 Pearson's correlation table for the CFR-Test and socio-economic factors for girls

	Categorical Free Recall Test	
	Pearson's r	Sign.
The education level of the father	-,083	,287
The father's occupation	,048	,372
Married or divorced or abroad	,083	,288
The number of siblings	-,218	,068
The order of birth	-,103	,242

Table 47 Pearson's correlation table for the CFR-Test and socio-economic factors for boys

	Categorical Free Recall Test	
	Pearson's r	Sign.
The education level of the father	,006	,483
The father's occupation	-,069	,312
Married or divorced or abroad	-,202	,074
The number of siblings	,246	**,038****
The order of birth	,185	,092

For the WST, the effect of the education level of the father disappeared for boys (see Table 49). However, it was still valid for girls (see Table 48). This may be due to the fact that girls are more "verbal" than boys, so they profit more from the influence of their fathers. In a literature review, Cicirelli (1978) also indicated that generally girls were better on verbal abilities than boys. Having found a positive relation between father's occupation and the WST, this result supports the view that fathers' occupation is related much more to girls' vocabularies than boys'.

Table 48 Pearson's correlation table for the WST and socio-economic factors for girls

	Word Span Test	
	Pearson's r	Sign.
The education level of the father	,191	,096
The father's occupation	,264	**,035****
Married or divorced or abroad	,054	,358
The number of siblings	,001	,498
The order of birth	-,029	,387

Table 49 Pearson's correlation table for the WST and socio-economic factors for boys

	Word Span Test	
	Pearson's r	Sign.
The education level of the father	,147	,147
The father's occupation	,179	,100
Married or divorced or abroad	-,033	,406
The number of siblings	,168	,115
The order of birth	-,195	,092

The WCST results were also affected by gender. Generally boys show the same pattern as the whole sample, i.e., their performance covaries with the marriage status of the parents and the number of siblings. However, this time the girls' performance was not under the influence of the marriage status of their parents and the number of their siblings. One possible reason for this situation might be that boys usually need a father figure in their family, i.e., a father maintains the "executive function" in a Turkish family. Since boys may need a role model for executive functions (father) and if this figure is away, their own executive functioning would be impaired. However, girls were not significantly affected by their parent's marriage status. Also, boys profited more from their siblings than girls did. As Irish (1964, as cited

Cicirelli, 1978, p. 365) reported "*Siblings may serve as role models for one another; particularly may the younger observe the other siblings of the same sex. They can serve as challengers and stimulators"* (p.282). Therefore, girls may have developed some strategies for distancing themselves from their siblings so that they may grow up more independently, however, they also do not profit from their siblings as much as the boys do. Moreover, some of the large-scale studies revealed that if the size of the family increases, the cognitive ability and achievement decreases (Cicirelli, 1978). Our results, however, provide counter-evidence to this negative relation: generally, the relation between the WCST and the number of siblings is positive, in particular for the boys.

Table 50 Pearson's correlation table for the WCST and socio-economic factors for girls

	Wisconsin Card Sorting Test	
	Pearson's r	Sign.
The education level of the father	,199	,088
The father's occupation	,098	,254
Married or divorced or abroad	-,155	,146
The number of siblings	,109	,231
The order of birth	-,047	,154

Table 51 Pearson's correlation table for the WCST and socio-economic factors for boys

	Wisconsin Card Sorting Test	
	Pearson's r	Sign.
The education level of the father	,073	,303
The father's occupation	-,182	,096
Married or divorced or abroad	-,339	**,007****
The number of siblings	,384	**,002****
The order of birth	,182	,376

For the LST, boys showed the same pattern as the general one, i.e., their performance covaried with the father's education level, number of siblings and birth order. However, again, the effects of number of siblings' and order of birth disappeared for girls. Both gender groups profited from the education level of their father. Moreover, when the numbers of sibling increases, boys profit from their siblings much more than girls, like in the WCST. Besides, boys' test results were better if they were the youngest of the family. Thus, the status of being "abi" (older brother) does not help much for this kind of test's performance. In a classical Turkish family, the oldest one is socially very much appreciated and the younger siblings have the lower status in hierarchy. At the same time, the intellectual capacities are inversely correlated to the birth order: the older boys perform less well than the younger boys. This seems to indicate that social rank and intellectual ability as measured in the LST are inversely related: as "abi" you don't have to do much, whereas as "youngest one" you have to make great intellectual efforts or you have to serve your older siblings more (for which you also need more intellectual resources) to get around in the family. This may differ between city to country-side, however.

However, the order of birth effect did not show up for girls. Since little girls are generally treated as little, cute, non-aggressive children, they might not want to behave competitively. Therefore this sisterly manner towards the other siblings may prevent triggering the order of birth effect for girls.

Table 52 Pearson's correlation table for the LST and socio-economic factors for girls

	Listening Span Test	
	Pearson's r	Sign.
The education level of the father	,305	**,018****
The father's occupation	,234	,055
Married or divorced or abroad	-,014	,463
The number of siblings	,228	,059
The order of birth	,067	,325

Table 53 Pearson's correlation table for the LST and socio-economic factors for boys

	Listening Span Test	
	Pearson's r	Sign.
The education level of the father	,283	**,020****
The father's occupation	,038	,394
Married or divorced or abroad	-,192	,085
The number of siblings	,379	**,003****
The order of birth	,273	**,024****

In total, for the gender effect, only three multiple regressions were found to be significant. One of them was the WCST for boys. As the correlation table indicates (see Table 51), the number of siblings – but not the marriage status of the parents– affected the results.

Table 54 Multiple regression results for the WCST and socio-economic factors for boys

Model		B	SE B	ß
Step 3	Constant	19,246	6,320	
	The number of siblings	2,345	,877	0,344*

Note R Square Change = .113 for Step 3, *p < .050.

Lastly, there were two significant multiple regressions results but just for boys and for the LST. These outcomes were also supported by the correlation tables.

Table 55 Multiple regression results for the LST and socio-economic factors for boys (order of birth)

Model		B	SE B	ß
Step 1	Constant	2,711	2,561	
	The education level of the father	2,796	1,270	0,332*
Step 3	Constant	1,378	2,965	
	The education level of the father	2,536	1,254	0,301*
	Birth order of the subject	,808	,387	0,274*

Note R Square Change = .090 for Step 1; R Square Change = .073 for Step 3, *p < .050

Table 56 Multiple regression results for the LST and socio-economic factors for boys (the number of siblings)

Model		B	SE B	ß
Step 1	Constant	2,711	2,561	
	The education level of the father	2,796	1,270	,332
Step 3	Constant	-1,605	3,130	
	The education level of the father	2,595	1,195	,308
	The number of siblings	1,351	,434	,393

Note R Square Change = .030 for Step 2; R Square Change = .148 for Step 3, *p < .050.

To sum up, The CFR-Test seemed not to be affected by all socio-economic factors and the gender factor. Only the boys were affected by the number of siblings which may be due to the fact that, in small villages, boys profit from their siblings and become more socialized. The WST was affected by both the education level of father and the occupation of the father and the gender factor. Since the WST just requires the memory capacity, the education level of father predictably affected the test results positively. However, it was just the girls who were affected by the occupation of the father. Because girls generally are found to be better on verbal abilities than boys, girls may take more advantage of their father's occupation than boys. For the WCST, girls did not display any socio-economic effect. What made the significant difference for the marriage status of the parents and the number of siblings factor for the whole group were the boys. Since the boys may need a father figure during this age period, those parental issues would affect the boys more than the girls. Lastly, the LST also showed the gender effect. Only the education level of father factor affected both girls and boys because the task requires both the storage and the manipulation of information and that can probably be supplied better by educated parents. Like in the CFR-Test and the WCST, boys were influenced by the number of siblings' factor. Additionally, this time the birth order positively influenced the boys. That is, the younger the boy, the better results for the LST. The reason for this might be that when a young boy joins the family, he might enter into a stimulating competition with his elder siblings. Thus, these competitive feelings might trigger the better outcomes in favor of the youngest boy.

To conclude, all these external influences related to socio-economic and family factors are probabilistic, they are not necessarily causal, or their causal direction is not always easy to determine. In the multiple regressions, however, we tried to work towards a better understanding of some of the causal links. Overall, boys and girls did not differ in intellectual capacity – they were on the same level in all memory tasks but they differed in the susceptibility towards such external influences. One important constraint of the study was the mother-related variables. They did not show up here because all mothers had the same status. However, they may influence the results. Further studies should be carried out in different parts of Turkey with a different socio-economic profile, namely with some urban population of Turkey.

CHAPTER 5

RESULTS (ADULTS)

5.1 The CFR-Test

5.1.1 Overall memory performance

According to the Kolmogorov-Smirnov Test and Shapiro-Wilk Test ($p < .05$), the data are non-normally distributed. Therefore non-parametric tests were used.

For the CFR-Test, the mean values, median, and standard deviation scores are shown in Table 57. Overall, from a 12-word list, adults remembered about 9 words.

Table 57 Descriptive statistics for the main task (adults)

	Mean	**S.D.**	**Median**
CFR-Test – absolute order of items	4,70	2,452	4,00
CFR-Test – absolute order of items&lists	7,45	2,685	6,00
CFR-Test – relative order of items	6,90	2,732	6,50
CFR-Test – relative order of items&lists	9,65	3,117	9,50
CFR-Test – total # of recalled items	9,10	1,861	9,00
CFR-Test – total # of recalled items&lists	11,95	2,139	12,00

We also separated the results of all tests according to the independent variable gender. For male and female students, the test statistics for the main task are shown in Figure 27. There was a significant difference between the results of the total number of remembered items (Z = -2.203, p = .031) and the total number of remembered items & lists (Z = -2.236, p = .025). That is, female students had M = 9.91, SD = 1.640 and U = 11 and male students had M = 8.11, SD = 1.691 and U = 9 on the total #of recalled items. Also, female students had M = 12.91, SD = 1.640 and U = 14 and male students had M = 10.78, SD = 2.167 and U = 12 on the total #of recalled items&lists. In Figure 27, the significant difference between two genders can be observed in terms of total #of recalled items and total #of recalled items&lists.

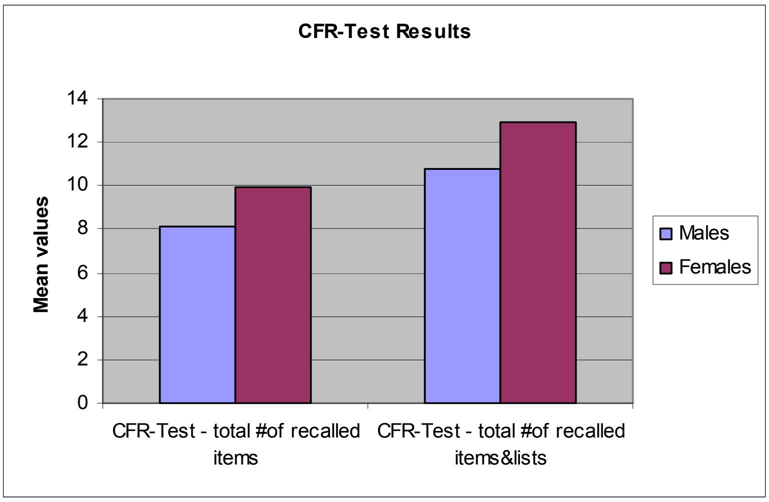


Figure 27 CFR-Test Results for the adult group of female and male students

5.1.2 Release from proactive interference

Overall, for the first category switch, there is no significant change in the number of remembered items between the last item of the first category and the first item of the second category. ($\chi 2(1)$ = 1,905, p = .150). For the 2nd category switch, that is, for the last item of the second category and the first item of the last category, the RPI is non-significant either ($\chi 2(1)$ = .960, p = .257).

The frequencies of the overall remembered items are shown in Figure 28 below. The insignificant levels of category shifts are indicated accordingly.

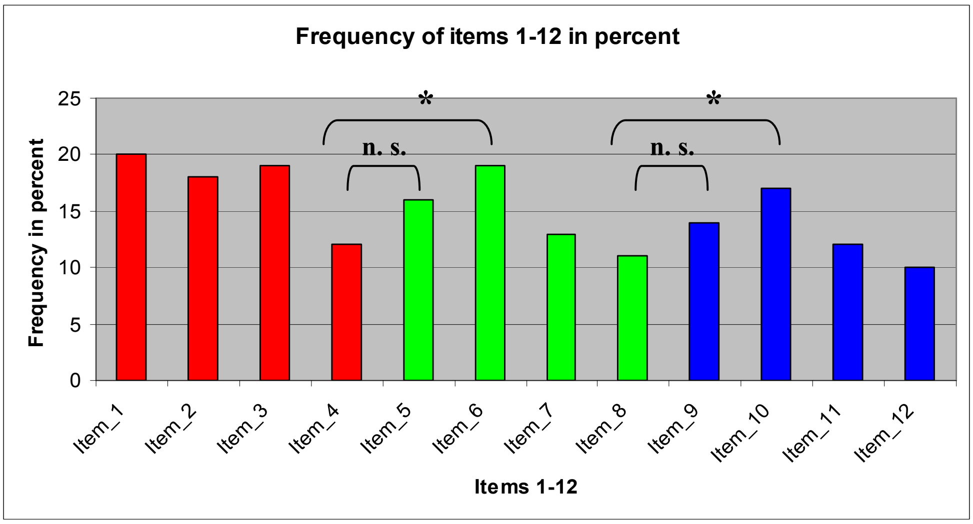


Figure 28 The absolute number of recalled items for the main task for the adult group

Interestingly, the RPI shows up at the 2nd item of the new category. Therefore, we tested the significance values between the last item of the 1st category and the 2nd item of the second category (item 4 and item 6) and the last item of the 2nd category and the 2nd item of the last category (item 8 and item 10). As expected, both these two shifts were significant (t (19) = -3.199, p = .005, and t (19) = -2.349, p = .030).

In general, the RPI effect could not be observed but a "delayed RPI effect" emerged when the 2nd items of the second and the third categories were taken into consideration.

5.1.2.1 Release from proactive interference and gender effect

There were no significant differences in the RPI between male and female students. For female students, the 1st change and the 2nd change were not significant (1st change, $\chi 2(1) = .917$, $p = .318$; 2nd change, $\chi 2(1) = 1{,}222$, $p = .293$). Similarly, for male students the two shifts were insignificant (1st change, $\chi 2(1) = .277$, $p = .500$; 2nd change, $\chi 2(1) = .234$, $p = .500$).

When the delayed shifts which occurred between c1_4 & c2_2 and c2_4 & c3_2 were examined, significant differences resulted between female and male subjects.

Namely, the first shift was significant for the female subjects (t (10) = -2.390, p = .038), and the second shift was significant for the male subjects (t (8) = -2.530, p = .035) according to the Paired Samples Test results. Therefore, it was the female subjects who showed the RPI effect for the first (delayed) category shift and it was the male subjects who showed the RPI effect for the second (delayed) category shift.

It can be concluded that there was no overall gender effect in the adult sample. However, gender effects occurred when the "delayed RPI effect" was taken into account. That is, the 1st shift was primarily because of the female subjects and the 2nd shift because of the male subjects.

5.1.3 False memories

In the adult group, there were just 3 subjects showing false memories and the numbers of false memories were 4. That is, nearly none of the subjects uttered incorrect words during the CFR-Test.

5.1.4 Analyzing the CFR-Test with ANOVA

5.1.4.1 Category, item, and gender interactions

The interactions between categories (3), items (4 for each category), and gender (2) were examined with ANOVA (for the mean table see Table 70 in Appendix D). As a between subjects factor gender was used and category & item were within subject factors. According to ANOVA results, category and item effects were significant (F (2, 20) = 9,748, p < .001; F (3, 20) = 8,184, p < .001). However, the interaction between item and gender, category and item, and item*category*gender were insignificant (F (3, 20) = .189, p = .903; F (6, 20) = .994, p = .433; F (6, 20) = 1.233, p = .295).

The adult subjects mainly showed both a category and item effect but did not display the interaction effect between the category and the item.

5.1.4.2 Gender effect for the main task

For the main task, two different measures were examined in order to asses the gender effect. The dependent variables were the overall number of remembered items and the absolute order of items & lists; and gender was used as between subject factor. Gender has no effect on the absolute order of items & lists ($F (1, 20) = 2.474$, $p = .133$). On the contrary, the total #of recalled items was significantly affected by gender ($F (1, 20) = 5.784$, $p = .027$). That is, female subjects (M = 9.91, SD = 1.640) remembered significantly more items than male subjects (M = 8.11, SD = 1.691).

In general, there was a significant gender difference in favor of the female subjects.

5.2 Additional memory tasks

The descriptive statistics for the additional tasks for the adults group are given in Table 58 below. Specifically, the results for the WST were 5 or higher for the adults. Also, all subjects have the highest level in the WCST, i.e. 6. For the Reading Span Test, the mean value for the level was 2.6 and most of the subjects' score were between 2 and 3.

Table 58 Descriptive statistics for the additional tasks (adults)

	Mean	S.D.	Median
Word Span Test	5,675	,335	5,50
Wisconsin Card Sorting Test	6,00	,000	6,00
Reading Span Test – level	2,60	,447	2,50
Reading Span Test – absolute # of correct items	13,40	3,185	13,00

5.2.1 The Reading Span Test

Since the Reading Span Test is a complex measure and newly developed, we also carried out some further statistics to understand this test better. Firstly, the correlation between the total reading time of the subject and the results of the test was examined. It was found that the total reading time had no relation with the Reading Span Test results, neither with the level nor with the absolute number of remembered items (see Table 59).

Table 59 Pearson's and Spearman's correlation table for the total reading time and the Reading Span Test

	Total reading time			
	Pearson's r	Sign.	Spearman's rho	Sign.
RST – level	-,021	,929	,102	,670
RST – absolute #of correct items	,035	,884	,131	,583

In the Reading Span Test, there were some subjects who tended to say "yes" for most of the sentences. We thought that there might be some effect of this "yes bias" on the results of the Reading Span Test. Therefore, we examined the correlation between the "yes bias of the subject" and the Reading Span Test scores. Again, there were no significant correlations (see Table 60).

Table 60 Pearson's and Spearman's correlation table for the yes bias of the subject and Reading Span Test scores

	Yes bias of the subject			
	Pearson's r	Sign.	Spearman's rho	Sign.
RST – level	,225	,340	,259	,270
RST – absolute #of correct items	,056	,814	,260	,269

Overall, it was found that both the total reading time and the yes bias of the subjects did not significantly affect the RST results.

Lastly, in the RST, the adults may have cheated in answering the questions faithfully because the performance of some subjects was very low. Therefore, the adult sample was separated in two groups. The criterion for successful performance was set at 85% (and above). Thus, we checked the differences between the successful sample (just 6 out of 20 subjects were above this criterion) and the non-successful sample (14 subjects) considering this criterion. We found no significant difference between these two samples. Further we omitted the non-successful subjects from the tasks and checked the differences between this new sample and the original sample. Again, we found no significant difference between these two samples. Consequently, it can be concluded that the results did not change because of the low performance of some unreliable subjects in the task.

5.2.3 Correlations between the tasks

The correlations between the additional tasks were also examined. Since all subjects had the same results for the WCST, it was excluded from the correlations. Therefore, only the correlation between the WST and the Reading Span Test was assessed. However, these two tests were not significantly correlated with each other in the adult group (see Table 61).

Table 61 Pearson's and Spearman correlation table for the WST and the Reading Span Test for the adult group

	Reading Span Test			
	Pearson's r	Sign.	Spearman's rho	Sign.
Word Span Test	,228	,333	,242	,305

To summarize, no significant correlation between the additional tasks was found.

5.2.5 Gender effect for the additional task

For the additional tasks, we found no gender effect at all (the WST, F (1, 20) = .313, *p* = .583; WCST, not applicable again; the Reading Span Test, F (1, 20) = .010, *p* = .923) (see also the descriptive statistics, Table 62). Thus the results revealed that there was no significant gender effect for the additional tasks.

Table 62 The results of the additional tasks for male and female students

	Gender of the student	Mean	S.E.
Word Span Test	man woman	5,722 5,636	,363 ,323
Wisconsin Card Sorting Test	man woman	6,00 6,00	,00 ,00
Reading Span Test - level	man woman	2,611 2,591	,486 ,437
Reading Span Test – absolute # of correct items	man woman	13,56 13,27	3,432 3,133

5.3 Multiple regressions

We analyzed two different multiple regressions for the main task. The first one was for the overall number of remembered items and the second is the absolute serial order of items & lists. The overall number of remembered items could not be predicted by any of the predictors (the WST, t (19) = 1.642, *p* = .120, r = .270; the Reading Span Test, t (19) = .118, *p* = .907, r = .088). Similarly, the absolute serial order of items & lists could not be predicted (the WST, t (19) = 1.050, *p* = .309, r = .288; the Reading Span Test, t (19) = .260, *p* = .789, r = .158). Generally, for the main task, neither the total #of recalled items nor the absolute order of items could be estimated by the additional tasks.

5.4 Graduation degree

We also ran a non-parametric test in order to examine whether there were any differences between the subjects who graduated from a Natural Science-related area vs. Social Science for all tests. According to the results of the Mann-Whitney Test, there was no difference between these groups (CFR-Test – total #of recalled items, Z = -.346, p = .739, CFR-Test – absolute order of items, Z = -.314, p = .796, the WST, Z = -.376, p = .739, the WCST, Z = .000, p = 1.000, the RST, Z = -.538, p = .631). Thus, the subjects did not differ in any of the tasks when the graduation degree effect was taken into consideration.

5.5 Discussion

Overall, the results of the main task were very good for the adult group. However, the RPI effect could not be observed, probably due to the fact that the memory capacity of the adults was relatively high. Although the exact RPI pattern could not be found, surprisingly a "delayed RPI" effect was observed that occurred on the 2nd item on the novel category and not on the first one. One possible explanation is that it may have taken time to realize the new category. In rapid presentation of the items the category shift may have been more fully realized by the adult subjects only when the second item of the new category occurred. Another possible explanation is that since the number of items in each category was relatively low (just 4 items) the category formation may not have been very strong at the point when the new category would open up. We also found gender difference, not for the standard RPI effect but for the delayed RPI effect. Namely, female subjects showed the first delayed shift and the male subjects the second delayed shift. Obviously, this delayed RPI effect and even more any possible gender difference need to be replicated in future studies. Besides, there were few records of false memories, indeed, i.e. just one subject uttered two items which were not in the list. Thus, false memories are almost inexistent in the adult group.

As the ANOVA showed, the category and the item effects were significant for the adults. That is, it did matter in which category and in which position they heard the items in the experiment. The category effect is caused by the second and third category being remembered less well as compared to the first one. The item effect is evidence for the PI and RPI effect. Also, we found a significant gender effect for the total # of recalled items. Here the female students were better than the male students. Generally, women tend to show better memory scores for capacity measures than men (see section 6.4 for possible explanations). However, for the additional tasks, no significant gender effect was found. Furthermore, the adult group performed well in the additional tasks except in the RST. Since the RST is a very difficult task to solve, the participants felt very distressed and even disappointed during the experiment. Thus, although the mean levels for this task indicated lower scores, in fact, the actual executive functioning level of the subjects was high enough for the task.

Furthermore, none of the results for the correlations and the multiple regressions were significant for the additional tests (see section 6.4 for possible explanations). Lastly, we found no effect of graduation degree (Natural Scientist vs. Social Scientist) which indicated that the adult sample showed similar result patterns for these kinds of tasks irrespective of the area of their academic studies.

CHAPTER 6

RESULTS (CHILDREN vs. ADULTS)

6.1 The CFR-Test

6.1.1 Overall memory performance

The differences between the 5th graders and the adults were compared since the oldest group in the children sample was the 5th graders. Thus, we wanted to see the exact contrasts between these two groups in order to conclude whether the development extends into the adulthood. Therefore, we tested the significance of the CFR-Test between the 5th graders and the adult group. According to the results, there were significant differences between these groups (see Table 65), i.e. adults scored significantly higher on all of the various measures for the CFR-Test (see the descriptives for the 5th graders and the adult group in Table 63-64)

Table 63 Descriptive statistics for the main task (5th graders)

	Mean	**S.D.**	**Median**
CFR-Test – absolute order of items	1,47	1,172	1,00
CFR-Test – absolute order of items&lists	3,05	1,747	3,00
CFR-Test – relative order of items	3,26	1,968	3,00
CFR-Test – relative order of items&lists	4,84	2,478	5,00
CFR-Test – total # of recalled items	5,63	1,461	6,00
CFR-Test – total # of recalled items&lists	8,37	1,921	9,00

Table 64 Descriptive statistics for the main task (adults)

	Mean	**S.D.**	**Median**
CFR-Test – absolute order of items	4,70	2,452	4,00
CFR-Test – absolute order of items&lists	7,45	2,685	6,00
CFR-Test – relative order of items	6,90	2,732	6,50
CFR-Test – relative order of items&lists	9,65	3,117	9,50
CFR-Test – total # of recalled items	9,10	1,861	9,00
CFR-Test – total # of recalled items&lists	11,95	2,139	12,00

Table 65 The test statistics of the Mann-Whitney Test for the main task for 5th graders and the adult group

	Z	**Asymp. Sig. (2-tailed)**
CFR-Test – absolute order of items	-4,328	,000
CFR-Test – absolute order of items&lists	-4,605	,000
CFR-Test – relative order of items	-3,853	,000
CFR-Test – relative order of items&lists	-4,082	,000
CFR-Test – total # of recalled items	-4,626	,000
CFR-Test – total # of recalled items&lists	-4,293	,000

6.1.2 Release from proactive interference

The release from proactive interference was overall significant for the children (first change, p = .022; second change, p = .045). However, considering the 5th graders, this effect was only instantiated partially (see section 4.1.2). For the adult group none of the two category changes were significant (first change, $\chi2(1)$ = 1,905, p = .150; second change, $\chi2(1)$ = .960, p = .257). However, we found significant differences between the last item of the previous category and the second item of the following category for both category switches in the adults. Such a "delayed" RPI effects were not generally observed in the 5th graders, however, there may be a hint of it in Figure 13 where the 5th graders also showed a higher recollection of the 2nd item of the last category when the mean number of remembered items for each position was assessed.

In section 4.1.2.2, we saw that there is no significant gender effect for the RPI. This is also valid for the adult group (for female students: 1st change, $\chi 2(1) = .917$, p = .318; 2nd change, $\chi 2(1) = 1,222$, p = .293; for male students: 1st change, $\chi 2(1) = .277$, p = .500; 2nd change, $\chi 2(1) = .234$, p = .500). However, there was one for the "delayed RPI effect" (see section 5.1.2.1).

Overall, both in the adult sample and in the 5th graders, the RPI effect could not be observed consistently.

6.1.3 False memories

According to the Kruskal-Wallis Test ($\chi 2(4) = 9.04$, $p = .06$), false memory decreases marginally over age for the school children. Also, for the adult group, we hardly observed false memories. Since there were just 4 false memories among 20 subjects, the adult data could not be properly analyzed statistically. However, it is clear that there is a meaningful difference between the children and the adult group on false memory. That is, false memories can be observed in young children however, they gradually disappear during development until for the adults they have vanished nearly completely. Brainerd et al. (2002) also supports this finding by stating that "False memories have typically been found to be more common during early childhood than during later childhood or adulthood" (p. 1363).

6.1.4 Analyzing the CFR-Test with ANOVA

6.1.4.1 Category and item interactions

In section 4.1.5.1, we found that for the children, the factors category and items are significant for the main task ($F(4, 101) = 10.029$, $p < .001$; $F(3, 101) = 6.686$, $p < .001$) (for the mean table see Table 71 in Appendix D). Similarly, for the adult group, category and item effects were significant ($F(2, 20) = 9,748$, $p < .001$; $F(3, 20)$ =

8,184, $p < .001$). The interaction between category and items was also significant ($F(6, 101) = 2.870$, $p = .009$) for the 5th graders. However, this interaction was insignificant for the adult group ($F(6, 20) = .994$, $p = .433$). Besides, the other interactions were insignificant for both the 5th graders and the adults.

The adult group was affected by both the category and the item effects like the 5th graders. However, they did not show the category*item interaction effect, i.e. it was not important that an item could be recalled differently when it was in the different categories.

6.1.4.1 Gender effect for the main task

The ANOVA results for the children showed that there is no significant effect of gender on the main task, the total #of recalled items ($F(1, 101) = 3.369$, $p = .070$) (see section 4.1.5.2). However, for the adult group, there was a significant effect of gender ($F(1, 20) = 5.784$, $p = .027$).

6.2 Additional memory tasks

The differences between the 5th graders and the adult group were also highly significant for the additional tasks except the Listening/Reading Span Test. Those two tests may not be directly comparable (for the descriptives see Table 66 and Table 67).

Specifically, adults remembered significantly more words in the WST (M = 5.675, SD = .335) than the 5th graders (M = 4.368, SD = .5973). It was also valid for the WCST (the adults, M = 6, SD = .00; the 5th graders, M = 4.47, SD = 1.020).

Table 66 Descriptive statistics for the additional tasks for the 5th graders

	Mean	S.D.	N
Word Span Test	4,368	,597	19
Wisconsin Card Sorting Test	4,47	1,020	19
Listening Span Test – level	2,737	,304	19
Listening Span Test – absolute # of correct items	12,053	2,345	19

Table 67 Descriptive statistics for the additional tasks for the adult group

	Mean	S.D.	N
Word Span Test	5,675	,335	20
Wisconsin Card Sorting Test	6,00	,000	20
Listening Span Test – level	2,60	,447	20
Listening Span Test – absolute # of correct items	13,40	3,185	20

Table 68 The test statistics of the Mann-Whitney Test for the main task between the adult group and the 5th graders

	Word Span Test	Wisconsin Card Sorting Test	Listening Span Test - level	Listening Span Test – absolute # of correct items
Z	-5,268	-5,744	-,958	-1,406
Asymp. Sig. (2-tailed)	,000	,000	,380	,166

6.2.4 Correlations between the tasks

The correlations between all of the additional tasks were highly significant for the children (see section 4.2.4). On the contrary, for the adult group, the two tests (WST and Reading Span test) were not correlated significantly (Pearson's r = .228, p = .333; Spearman's rho = .242, p = .305). Thus, in opposition to the 5th graders, none of the tasks was significantly correlated for the adult group.

6.2.5 Gender effect for the additional task

We found no significant gender effect for the additional tasks in the children's data (see section 4.2.5). For the adult group, also, there was no significant gender effect for the other tests (the WST, F (1, 20) = .313, p = .583; the WCST, not applicable; the RST, F (1, 20) = .010, p = .923). Therefore, no gender effect was observed for these two groups.

6.3 Multiple regressions

For the children group, we found that the WST for the children could predict the overall # of remembered items (t (100) = 3.434, p = .001) and the LST could (marginally) predict serial order (t (100) = 1.967, p = .053). However, for the adult group, the main task could not be predicted by any of the additional tasks (the WST, t (19) = 1.642, p = .120, r = .270; the Reading Span Test, t (19) = .118, p = .907, r = .088). So, overall, for the adult sample, the main task could not be estimated by any of the additional tasks.

6.4 Discussion

Overall, the adult group was better on all tasks except for the LST/RST. However, for the RPI effect, we again could not exactly obtain the predicted pattern for the adults. That is, adults did not show the standard RPI effect on the first item of the two new categories, respectively. Since the general pattern found for the children in all grades was instable anyway, it was not unusual to obtain inconsistent results for the adult sample as well. We explained the missing RPI effect with the overall high memory performance of the adults which would level the differences between the various positions. However, the delayed shifts described in section 5.1.2 deserve some closer attention. The fact that a significant recovery of memory occurred not on the first item of the new category but on the second may be due to the following reasons: Firstly, for the adults, the time to realize that there was a category shift took longer. When the subject hears an item from a different category after having listened to the items from the previous category, s/he might not realize that there is a shift in the categories. Then, as soon as the second item of the different category comes into play, the change in the categories is confirmed with more certainty. Therefore, the subject may only then have become aware of the shift. There may be evidence for such a delayed RPI effect in the second category shift of the 5th graders as well, who may behave more similar to the adults than the younger ones. In our experiment, categorical cohesion and categorical distinction were very different than the ones in the literature because we had two category shifts. In a standard RPI task, where 3 sets of 4 items from the same category (e.g. fruits) are used which are then followed by a set of items from a different category (e.g., animals), the RPI effect is stronger than it was in our study (see section 2.3). The reason why the standard RPI effect was not found in the adults but rather a delayed RPI effect arose in the adults and possibly also in the 5th graders, could be the insufficient strength of the build-up of the PI effect in the current study. Another reason might also be that the presentation time of the items was relatively fast. Obviously, these initial findings should be followed up by further studies.

For the main task, the category effect indicates a primacy effect. That is, the first category is always recalled better than the following categories. The item effect also shows the build-up of PI since in a category the first items are remembered best while the last items are recalled worst. Lastly, the category*item interaction reveals that it makes a difference for an item to be in a particular category (position) and to be in the different item position in a category. Since the children are sensitive to both the item and the category positions while they are developing, they did show this significant interaction. However, for the adults, this was not the case anymore. Since there does not seem to be any development for the adults in terms of memory capacity anymore, they were not sensitive to these interactions. Given that the experiment was not challenging enough for the adult, their overall recall level was very good in the main task. Thus, the categories were recalled nearly equally well. They benefited from the (delayed) RPI effect by means of their good memory and categorization abilities and showed no interaction effect.

The other issue was the difference between the female subjects and the male subjects for the adult group. The female subjects were better than male subjects in the total #

of recalled items. However, it was the (not significantly) reverse for the child sample in the higher grades (see section 4.1.1). There might be two reasons for this change: The first one could be that girls might develop between the age range of 12-20 years. The second reason could be due to a social effect. One possible reason might be that girls are given less freedom to express themselves in their families than the boys in a classical Turkish family, as typical for the rural area where the children's experiment has been carried out. So the executive memory abilities of the girls may only develop later in time than those of boys. Girls (of rural areas) that continue with their school education and enter university later on will constitute a highly selected sample of those initial cohorts. Those girls become less restricted than they were when they were younger. That is, either they are those who had already strong cognitive abilities and/or they become more self-confident, more verbal, and more developed in terms of their cognitive abilities through their continued studies. Obviously, the differences between subjects in rural and urban areas have to be taken into consideration. Since the adult experiments were conducted in an urban area (Ankara) and the children's experiments in a rural area (Yozgat), there might actually be great differences between these samples. Because in a rural area boys are treated as more prestigious than girls, this could affect the results in favor of the boys.

The results of the WST increased for the adult group as compared to the 5th graders, as expected. For the WCST, all adults attained the highest level since the test is generally used in clinical samples. Thus, the results confirm that our sample included clinically normal adults. Lastly, the RST was a relatively hard task for the adults. In the same way, the LST was a hard experiment for the children. Therefore, no significant difference between the two samples was found. Note that the LST could not be used in the adult group since it would have been too easy and they would not have taken it seriously.

Lastly, the results of the correlations and multiple regressions revealed that there was no significant correlation between the additional tasks and the main task in the adult sample and hence the additional tasks could not predict the CFR-Task. This is probably due to the fact that the memory capacity of the adults does not develop further. If there is no development, a considerable part of the variation may be lost and there should be no more significant correlations. This means that most of the predictive power of the additional memory tasks on the main task in the children can be attributed to their common development. Another factor that may have made the correlations in the adult sample insignificant is related to different aspects of memory capacity in the WST and the RST. Since it seemed that both tasks measured different aspects of working memory (the WST measuring overall short-term verbal memory capacity, the RST measuring complex working memory) the resources shared by these tasks should be minimal. However, for the children all tasks were quite related to each other according to the correlations (see section 4.2.4). Because the memory abilities of children are still developing, the resources for the various tasks could be still shared more than those of the adults, that is, the memory system of the children is still differentiating.

We suggest that these findings should be followed up by either a longitudinal study investigating the same children when they are grown up or by a cross-sectional comparative study in which subjects from an urban area are compared with our subjects of a rural area.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSIONS

In this study, we aimed to examine the RPI phenomena and its various relations with other working memory components, namely phonological, complex, and executive working memory functions. Since we discussed the findings already in depth in the various discussion sections in the result part of this thesis, this chapter will highlight the main findings and draw conclusions from it.

To start with the main task, the CFR-Test, we observed a significant development throughout childhood. Thus, our 1^{st} hypothesis was confirmed which indicates that the overall memory span increases with age. However, our 2^{nd} and 3^{rd} hypotheses were not confirmed. Neither the categorization ability nor the RPI effect increased with the age of the children. There was no significant difference between younger and older children in this respect. With respect to the PI effect, we had no directed hypothesis. It turned out that the PI effect did not change either. It was stable throughout childhood (in line with findings cited in Kail (2002)). However, in itself the RPI was unstable, that is, it was not reliably present or absent throughout the children's development. Since the categorization ability, the PI and the RPI effects were present even in 1^{st} graders, it did not develop during the childhood. One possible reason is that by the second years of life the categorization ability starts emerging in infants (Younger and Fearing, 1999) and this ability increases during the childhood so the genuine development of categorization could appear even earlier than school-age, i.e., before 6 years of age. Thus, we could not observe the significant development after primary school-age. Also, the re-organization of the memory that seemed to take place in the 3rd grade may have spoiled both the PI and the RPI pattern that was present for the 1st and the 2nd graders in the sense that an overall flat memory curve would prevent any RPI effect to surface. Besides, the shortness of the list was a factor for the 4th and the 5th graders because their relatively higher memory capacity might not have allowed these effects to occur.

The 4th hypothesis, that is, the performance on the main task and on the RPI effect of the adults should be higher than that of children, was confirmed partially. The adult group showed higher performance in all tasks except the RST which was relatively hard. However, it was not confirmed with respect to the RPI pattern since the RPI patterns produced by both children and adults were very unstable. Thus, it cannot be concluded that in the adult sample the RPI pattern was more pronounced than it was in the child sample. Like for the older children, it was due to the fact that the memory capacity of the adults was relatively high so the RPI pattern could not be clearly observed in the adult group.

We also observed a strong developmental progression in the additional memory tasks, i.e. the WST, the WCST, and the LST. That means, our 5th hypothesis was also confirmed. The increase was linear for the WST and the WCST but it was step-wise for the LST. These developmental patterns confirm the findings of Gathercole (1999) who also found a steady increase for simple working memory tasks across a comparable age range but a slower and stepwise increase for complex working memory. These additional tasks were used as predictors of the main task, in order to find out the underlying structure of the CFR test. All tasks were highly predictive of the main task for children, thus confirming the last hypothesis of the study, but not for the adults. This result reveals that the good predictability in childhood was due to the common development of the tasks, whereas the predictability was low in the adults since they did not develop further and had uniformly attained high levels of performance. Especially, the WST became highly significant for the prediction of the children's overall memory capacity in the main task when all tasks were entered at once in a multiple regressions. This is because the CFR-Test generally measures the overall working memory capacity like the WST does. Then, the WST absorbs all the variance from the other additional tasks since there was a high correlation between the WST and the other two tasks, the WCST and the LST. However, when the criterion was changed, e.g. the performance on the correct serial order of items had to be predicted, it was the LST that was most predictive among all tasks. From this finding we can conclude that serial recall of items involves executive and complex WM functions to a higher degree. Executive and complex WM functions are particularly measured by the LST.

The socio-economic factors were not related to the results of the main task but to the additional tasks. We looked at three general factors regarding the socio-economic dynamics of a family, namely, the education level/occupation of the father, the family/marriage status of the parents, and the number of siblings/ the birth order. Firstly, the education level of the father covaried positively with both the WST and the LST which points to its relation with the vocabulary of the children. The occupation of the father only correlated with the WST which shows that education, not occupation, is more connected with children's memory development. The family status (father being abroad) had a negative relation with the WCST. One possible reason is that the WCST measures one of the executive functions which are affected by the physical presence of the father at home. The child may need a father figure which fosters the development of her/his own executive functions. For both the WCST and the LST, the number of sibling factor was significant. That is, the more the siblings you have the more successful you are in complex WM functions. Thus, the communication between the siblings might trigger the use of executive functions more in families with many siblings than in families with fewer siblings. Lastly, the birth order only correlated with the LST. Namely, the youngest of the family were most likely to attain the highest scores among the siblings. This might be due to the fact that the youngest one in a family might feel inferior to her/his elder siblings and constantly has to cope in order to keep up with them. Therefore, s/he would become more competitive which helps her/his executive functions to develop.

Overall, in most of the results no gender effect was observed. For the CFR-Test and the additional tasks, there was no significant difference between girls and boys for the child sample. However, in the adult sample, the female subjects were significantly better as compared to the male subjects in terms of the total # of recalled items but not for the additional tasks. This result contrasted with the finding that the girls were slightly inferior to the boys by the end of the 5^{th} grade. Thus, the young female students might have experienced a relatively late development in adolescence. However, it can also be that the child and the adult sample differed in other respects, e.g., the child sample was drawn from a rural area whereas the adult sample was drawn from an urban area. For the CFR-Test, the developmental pattern was slightly different for the boys and girls: the boys developed significantly in the total number of recalled items, whereas the girls in the number of remembered items in correct serial order. That is, the boys had slightly higher scores in seriality on which they did not improve too much and slightly lower scores in overall memory capacity on which they improved a lot. The reverse pattern was found for girls. For the RPI effect, overall, girls and boys did not differ in their results, neither in the child nor in the adult sample. Lastly, the socio-economic factor also displayed gender effects. For example, the occupation of the father only correlated positively with the WST because of the girls but not the boys. Therefore, it seems that girls' language working memory development was stronger related to their fathers' occupation than the boys'. Conversely, the negative correlation of the family status of the father being abroad with the WCST was due to the boys only. That is, the absence of the father figure had a stronger negative impact on boys than on girls. For both the WCST and the LST, the number of sibling factor was again only related to the boys. In this sample, girls did not seem not having an advantage of having more siblings as opposed to boys. This finding may correspond to the more favorable status of boys rather than girls in a rural area in which these experiments were conducted. Thus, boys profit from their siblings more than girls. Lastly, for the LST, it was again the boys for who birth order was significant. Boys profited from being the youngest in their family but not girls. The "advantage of being a boy" again showed up in this situation

In the following section, I would like to address some weaknesses of this study. Firstly, the study was conducted in a rural area for which there exist no comparative studies in the developmental area. Therefore it is hard to draw any strong and final conclusions from our sample. Still, it was an advantage to conduct the experiments in Yozgat. Our study which assessed a multitude of very general memory and cognitive processes points to a blind spot in the literature and explicitly deals with the peculiarities of a rural sample. Thus, the current study should be compared with other studies in urban Turkish areas in the future in order to find out whether the cognitive memory processes are stable everywhere irrespective of being carried out in a rural or in an urban area. The cognitive and developmental resilience of memory processes could thereby be addressed.

Secondly, there was no possible direct comparison for the LST and the RST results since in the literature there are only English studies on these experiments. Because the LST was adapted for Turkish and the RST was newly developed, future studies following this one can make use of both tests in order to compare the similarities and the differences between their results and ours.

Thirdly, the PI and the RPI effects could not be observed clearly since the stimulus list might have been too short. That is, the task was slightly too easy for the older children and the adults. However, if it had been longer, then it would have been too hard for the younger children. Instead, the same stimulus list was used both for the younger and the older subjects as well as for the adults. Although the list did not always allow the PI and the RPI effect to manifest itself clearly, the fact that there were two category shifts which on and off showed the RPI effect allowed us to observe the effect of the categorization ability on the organization of working memory in general. In this constellation, surprisingly, a novel effect emerged in adults (and maybe also partially for the 5th graders): the "delayed RPI effect". It seems that if the various categories comprise only few items, the category shift may be detected only belatedly. This suggests that the strength of category formation and the point at which a novel category is identified depends on the number of the items in the category. This new finding should be studied further by changing the number of categories, the number of the items in the categories, and the duration of the stimuli in order to systematically observe different patterns of PI and (delayed) RPI effects.

The current study was carried out with both children (in a rural area) and adults (in an urban area). It was a study with a wide age range, drawn from a big sample in Turkey. The main task, the RPI experiment, was different from the other RPI tasks in the literature since we had two category shifts. The additional tasks, except the WCST, were newly developed. The WST was taken from Bayramoglu and Hohenberger (2007) and further improved in order to study the phonological WM of children. We also used the children's version of the WCST since not the original version but the Modified WCST is recommended for children. The aim of the test was to measure both the categorization ability of children and their executive functions. There were two experiments regarding complex working memory: the LST (for children) and the RST (for adults). The original study of Pickering and Gathercole (2001), the LST, was adapted for Turkish and the RST was inspired by Saito and Miyake (2004) and developed by the author.

The LST was translated from English to Turkish and adapted to the language rules of Turkish. Future studies in this area may both use the test and compare their results with ours. Besides, the RST can also be developed further for future studies. Different versions could be obtained by changing the level of difficulty of the sentences, by conducting the experiment with an experimenter instead of as a self-paced study, and adding a point for each sentence that is answered correctly.

To conclude, with this study, we want to contribute to the knowledge on the development of the human working memory system. Examining and comparing both the results from the children and the adults by means of the CFR-Test and the additional working memory tests could help us improve our knowledge in this central area of cognitive science.

REFERENCES

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and Visuospatial Short-Term and Working Memory in Children: Are They Separable? *Child Development,* 77(6), 1698-1716.

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A.-M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology,* 87(2), 85-106.

Anastasi, A. (1956). Intelligence and family size. *Psychological Bulletin,* 53(3), 187-209.

Anderson, M.C., Neely J.H. (1996). Interference and inhibition in memory retrieval. In: Memory. Handbook of perception and cognition, 2nd ed. (Bjork EL, Bjork RA, eds), pp 237–313. San Diego, CA: Academic Press.

Archibald, L. M. D., & Alloway, T. P. (2008). Comparing language profiles: children with specific language impairment and developmental coordination disorder. *International Journal of Language & Communication Disorders,* 43(2), 165 - 180.

Ashby, F. G., & Maddox, W. T. (2005). Human Category Learning. *Annual Review of Psychology,* 56(1), 149-178.

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. *Trends in Cognitive Sciences,* 9(2), 83-89.

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

Baddeley, A. D. (1996). Exploring the Central Executive. *The Quarterly Journal of Experimental Psychology Section A,* 49(1), 5 - 28.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences,* 4(11), 417-423.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 829-839.

Baddeley, A. D., Hitch, G., & Gordon, H. B. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. Volume 8, pp. 47-89): Academic Press.

Badre, D., & Wagner, A. D. (2005). Frontal Lobe Mechanisms that Resolve Proactive Interference. *Cerebral Cortex,* 15(12), 2003-2012.

Barbara A. Younger, D. D. F. (1999). Parsing Items into Separate Categories: Developmental Change in Infant Categorization. *Child Development,* 70(2), 291-303.

Bayley, N. (1965). Comparisons of Mental and Motor Test Scores for Ages 1-15 Months by Sex, Birth Order, Race, Geographical Location, and Education of Parents. *Child Development,* 36(2), 379-411.

Bayramoglu, M., & Hohenberger, A. (2007). *The development of Theory of Mind in Turkish preschoolers*. Paper presented at the 13th European Conference on Developmental Psychology (ECDP).

Bjerkedal, T., Kristensen, P., Skjeret, G. A., & Brevik, J. I. (2007). Intelligence test scores and birth order among young Norwegian men (conscripts) analyzed within and between families. *Intelligence,* 35(5), 503-514.

Blake, J. (1981). Family Size and the Quality of Children. *Demography,* 18(4), 421-442.

Boomsma, D. I., van Beijsterveld, T. C. E. M., Beem, A. L., Hoekstra, R. A., Polderman, T. J. C., & Bartels, M. Intelligence and birth order in boys and girls. *Intelligence,* In Press, Corrected Proof.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are Young Children Susceptible to the False Memory Illusion? *Child Development,* 73(5), 1363-1377.

Breland, H. M. (1974). Birth Order, Family Configuration, and Verbal Achievement. *Child Development,* 45(4), 1011-1019.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of Experimental Psychology,* 10, 12-21.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology,* 17(2), 243-250.

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development,* 72(4), 1032-1053.

Cermak, L. S., Sagotsky, G., & Moshier, C. (1972). Development of the ability to encode within evaluative dimensions. *Journal of Experimental Child Psychology,* 13(1), 210-219.

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified Wisconsin Card Sorting Test (MCST, MWCST): Normative Data in Children 4–13 Years Old, According to Classical and New Types of Scoring. *The Clinical Neuropsychologist,* 21(3), 456 - 478.

Cicirelli, V. G. (1977). Children's school grades and sibling structure. *Psychological Reports,* 41, 1055-1058.

Cicirelli, V. G. (1978). The Relationship of Sibling Structure to Intellectual Abilities and Achievement. *Review of Educational Research,* 48(3), 365-379.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior,* 19(4), 450-466.

Darling, S., & Valentine, T. (2005). The categorical structure of semantic memory for famous people: a new approach using release from proactive interference. *Cognition,* 96(1), 35-65.

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review,* 12(1), 45-75.

Dempster, F. N., & Corkill, A. J. (1999). Interference and inhibition in cognition and behavior: Unifying themes for educational psychology. *Educational Psychology Review,* 11, 1-88.

Derwing, B. L., & de Almeida, R. G. (2004). Non-chronometric experiments in linguistics. In D. E. (Ed.) (Ed.), *Experimental and Quantitative Linguistics*.

Douglas, J. D., & Corsale, K. (1977). The Effects of Mode and Rate of Presentation on Evaluative Encoding in Children's Memory. *Child Development,* 48(1), 46-50.

Downey, D. B. (2001). Number of Siblings and Intellectual Development: The Resource Dilution Explanation. *American Psychologist,* 56(6-7), 497-504.

Du, Y., Xiao, Z., Song, Y., Fan, S., Wu, R., & Zhang, J. X. (2008). An electrophysiological signature for proactive interference resolution in working memory. *International Journal of Psychophysiology,* 69(2), 107-111.

Dukas, R., & Waser, N. M. (1994). Categorization of food types enhances foraging performance of bumblebees. *Animal Behaviour,* 48(5), 1001-1006.

Engle, R. W., Kane, M. J., & Brian, H. R. (2003). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. In *Psychology of Learning and Motivation* (Vol. Volume 44, pp. 145-199): Academic Press.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. M. P. S. (Eds.) (Ed.), *Models of working memory* (pp. 102-134). Cambridge: Cambridge University Press.

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior,* 11(6), 778-783.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences,* 3(11), 410-419.

Gathercole, S. E., & Alloway, T. P. (2004). Working memory and classroom learning. *Professional Association for Teachers of Students with Specific Learning Difficulties,* 17, 2-12.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology,* 40(2), 177-190.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology,* 18(1), 1-16.

Geis, M. F. (1975). A developmental study of category, rhyme, and color encoding in a release-from-proactive-interference paradigm. Emory University.

Glass, D. C., Neulinger, J., & Brim, O. G., Jr. (1974). Birth Order, Verbal Intelligence, and Educational Aspiration. *Child Development,* 45(3), 807-811.

Halford, G. S., Maybery, M. T., O'Hare, A. W., & Grant, P. (1994). The Development of Memory and Processing Capacity. *Child Development,* 65(5), 1338-1356.

Hamilton, A. C., & Martin, R. C. (2007). Proactive Interference in a Semantic Short-Term Memory Deficit: Role of Semantic and Phonological Relatedness. *Cortex,* 43(1), 112-123.

Hasher, L., Chung, C., May, C. P., & Foong, N. (2002). Age, Time of Testing, and Proactive Interference. *Canadian Journal of Experimental Psychology,* 56(3), 200-207.

Hasher, L., Zacks, R. T., & Gordon, H. B. (1988). Working Memory, Comprehension, and Aging: A Review and a New View. In *Psychology of Learning and Motivation* (Vol. Volume 22, pp. 193-225): Academic Press.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D. A., Martin, S., & Stuart, G. (1997). Word-Frequency Effects on Short-Term Memory Tasks: Evidence for a Redintegration Process in Immediate Serial Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 23(5), 1217-1232.

Inoue, T., Hasegawa, T., Takara, S., Lukáts, B., Mizuno, M., & Aou, S. (2008). Categorization of biologically significant objects, food and gender, in rhesus monkeys: I. Behavioral study. *Neuroscience Research,* 61(1), 70-78.

Irish, D. P. (1964). Sibling Interaction: A Neglected Aspect in Family Life Research. *Social Forces,* 42(3), 279-288.

Jaeger, J. J. (2005). *Kids' Slips: What Young Children's Slips of the Tongue Reveal about Language Development*: Lawrence Erlbaum Associates. 10 Industrial Avenue, Mahwah, NJ, 07430. Tel: 800-926-6579; Fax: 201-760-3735; e-mail: orders@erlbaum.com; Web site: http://www.erlbaum.com.

Jonides, J., & Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience,* 139(1), 181-193.

Just, M. A., & Carpenter, P. A. (1999). A capacity theory for comprehension. *Psychological Review,* 99, 122–149.

Kail, R. (2002). Developmental Change in Proactive Interference. *Child Development,* 73(6), 1703-1714.

Kail, R., & Hall, L. K. (2001). Distinguishing short-term memory from working memory. *Memory & Cognition,* 29, 1-9.

Kail, R. V., & Levine, L. E. (1976). Encoding processes and sex-role preferences. *Journal of Experimental Child Psychology,* 21(2), 256-263.

Kail, R. V., & Schroll, J. T. (1974). Evaluative and taxonomic encoding in children's memory. *Journal of Experimental Child Psychology,* 18(3), 426-437.

Kee, D. W., Gregory-Domingue, A., Rice, K., & Tone, K. (2005). A release from proactive interference analysis of gender schema encoding for occupations in adults and children. *Learning and Individual Differences,* 15(3), 203-211.

Kee, D. W., & Helfend, L. G. (1983). Developmental differences in the encoding of spatial-orientation information. *Bulletin of the Psychonomic Society,* 21, 381–383.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior,* 1(3), 153-161.

Kluender, K. R., Diehl, R. L., & Killeen, P. R. (1987). Japanese Quail can form phonetic categories. *Science,* 237, 1195-1197.

Lewandowsky, S., & Farrell, S. (2008). Phonological similarity in serial recall: Constraints on theories of memory. *Journal of Memory and Language,* 58(2), 429-448.

Marques, J. F. (2000). The 'living things' impairment and the nature of semantic memory organisation: An experimental study using PI-release and semantic cues. *Cognitive Neuropsychology,* 17, 683–707.

McGeoch, J. A. (1942). *The psychology of human learning*. New York: Longmans, Green.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review,* 63(2), 81-97.

Nakayama, S. Y., & Kee, D. W. (1980). Automatic encoding of superordinate and subordinate taxonomic categories in different populations. *Journal of Educational Psychology,* 72(3), 386-393.

Nee, D. E., Jonides, J., & Berman, M. G. (2007). Neural mechanisms of proactive interference-resolution. *NeuroImage,* 38(4), 740-751.

Nimmo, L. M., & Roodenrys, S. (2004). Investigating the phonological similarity effect: Syllable structure and the position of common phonemes. *Journal of Memory and Language,* 50(3), 245-258.

Passler, M. A., Isaac, W., & Hynd, G. W. (1985). Neuropsychological development of behavior attributed to frontal lobe functioning in children. *Developmental Neuropsychology,* 1, 349–370.

Paul C. Quinn. (2002). Category Representation in Young Infants. *Current Directions in Psychological Science,* 11(2), 66-70.

Pender, N. J. (1969) A developmental study of conceptual, Semantic Differential, and acoustical dimensions as encoding categories in short-term memory. Final report of Project No. 9-E-070, U.S. Department of Health, Education, and Welfare, Northwestern University, Washington, D.C.: Government Printing Office.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology,* 58(3), 193-198.

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. London: Psychological Corp.

Repovs, G. & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience,* 139(1), 5-21.

Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology,* 12, 105–126.

Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in reading span performance. *Journal of Memory and Language,* 50(4), 425-443.

Schooler, C. (1973). Birth order effects: A reply to Breland. *Psychological Bulletin,* 80(3), 213-214.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory & Cognition,* 21, 168–175.

Stratta, P., Daneluzzo, E., Prosperini, P., Bustini, M., Mattei, P., & Rossi, A. (1997). Is wisconsin card sorting test performance related to `working memory' capacity? *Schizophrenia Research,* 27(1), 11-19.

Sulloway, F. J. (2007). Psychology: Birth Order and Intelligence. *Science,* 316(5832), 1711-1712.

Thorn, A. S. C., Gathercole, S. E., & Frankish, C. R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert's (1993) multinomial processing tree model. *Cognitive Psychology,* 50(2), 133-158.

Tillman, C. M., Nyberg, L., & Bohlin, G. (2008). Working memory components and intelligence in children. *Intelligence,* 36(5), 394-402.

Tyrrell, D. J., Pressman, M., Cunningham, T., Steele, G. M., & Thaller, K. (1981). Input and distractor modality effects upon the release from proactive interference in children. *Journal of Genetic Psychology,* 139, 205–220.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review,* 64(1), 49-60.

Vuontela, V., Steenari, M.-R., Carlson, S., Koivisto, J., Fjallberg, M., & Aronen, E. T. (2003). Audiospatial and Visuospatial Working Memory in 6-13 Year Old School Children. *Learning & Memory,* 10(1), 74 - 81.

Whitney, P., Arnett, P. A., Driver, A., & Budd, D. (2001). Measuring Central Executive Functioning: What's in a Reading Span? *Brain and Cognition,* 45(1), 1-14.
Wickens, D. D. (1970). Encoding categories of words: An empirical approach to meaning. *Psychological Review,* 77(1), 1-15.

Wickens, D. D. (1972). Characteristics of word encoding. In A. M. E. M. (Eds.) (Ed.), *Coding processes in human memory*. Washington: DC: Winston.

Winston D. G. (1999). Lexical Neighborhoods and Release from Proactive Interference: A First Report .*Research on Spoken Language Processing.* Progress Report No. 23. Indiana University

Younger, B. A., & Fearing, D. D. (1999). Parsing Items into Separate Categories: Developmental Change in Infant Categorization. *Child Development,* 70(2), 291-303.

Zajonc, R. B. (1976). Family Configuration and Intelligence. *Science,* 192(4236), 227-236.

Zajonc, R. B. (2001). The Family Dynamics of Intellectual Development. *American Psychologist,* 56(6-7), 490-496.

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early Development of Executive Function: A Problem-Solving Framework. *Review of General Psychology,* 1(2), 198-226.

APPENDICES

APPENDIX A EXPERIMENTS

Experiment 1 (The Word Span Test)

<u>2'LİK SETLER</u>

Köşk	-	Muz
Pil	-	Üst
Buz	-	Dört

<u>3'LÜK SETLER</u>

Göl	-	Saç	-	Tuz
Sev	-	Kürk	-	Bel
Kir	-	Ut	-	Pas

<u>4'LÜK SETLER</u>

Kaş	-	Sos	-	Göç	-	Yat
Cam	-	But	-	Sal	-	Köy
Zar	-	Kuş	-	Tüm	-	Can

5’LİK SETLER

Suç - Kek - Böl - Top - Zam

Bal - Kurt - As - Tat - Çöp

Ot - Son - Türk - Seç - Kol

6’LIK SETLER

Hak - Sus - Tek - Mum - Dip - Kar

Kes - Bin - Ter - Aşk - Yut - Sel

Tren - Kel - Söz - An - Koy - Tez

7’LİK SETLER

Ak - Top - Su - Alt - Bey - Bol - Mart

Tel - Poz - At - Bil - Yok - Fes - Tür

Kış - Ver - Han - Bot - Yıl - Post - Kül

8’LİK SETLER

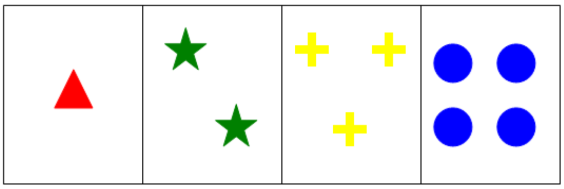
Tam - Bak - Uç - Göz - Hal - Boş - Ek - Yurt

Üç - Kas - Al - Mülk - Bir - Tut - Dil - Kum

Bul - Pek - On - Fal - Var - El - Ses - Genç

Experiment 2 (Wisconsin Card Sorting Test)

Four stimulus cards



Experiment 3 (Listening Span Test)

Deneme seti

1. Çocuklar okula gider.
2. Balıklar havada yaşar.
3. Ağaçlar dans eder.

2'LİK SETLER

1

1. Biber acıdır.
2. Kediler okulda çalışır.

2

1. Filler çok küçüktür.
2. Ayakkabı ayağa giyilir.

3

1. İnsanlar saçlıdır.
2. Çicekler fare kovalar.

4

1. Ayılar araba sürer.
2. Havuçlar turuncudur.

5

1. Gece karanlıktır.
2. Portakallar suda yaşar.

6

1. Ateş sıcaktır.
2. Balıklar konuşur.

3’LÜK SETLER

1

1. Otobüslerle tatile gideriz.
2. Toplar karedir.
3. Öğretmenler ağaçta yetişir.

2

1. Muzlar bisiklete biner.
2. Elimiz beş parmaklıdır.
3. Soğan acıdır.

3

1. Otobüsler oyuncakla oynar.
2. Kuşlar kanatlıdır.
3. Elmalar ağaçta yetişir.

4

4. Piyanolar müzik çalar.
5. Kardeşlerimiz kuyrukludur.
6. Burnumuzla görürüz.

5

4. Ayağımız çenelidir.
5. Güneş sıcaktır.
6. Taşlar serttir.

6

4. Kaşıklarla yazı yazarız.
5. Limon sarıdır.
6. Köpekler kedileri kovalar.

4'LÜK SETLER

1

1. Zürafalar uzun boyludur.
2. Çiçekler pasta sever.
3. Portakallar kulaklıdır.
4. Öğretmenler okulda çalışır.

2

1. Otobüsler konuşur.
2. Bankalardan para çekeriz.
3. Kışlar sıcaktır.
4. Pastalar tatlıdır.

3

1. Gökyüzü kırmızıdır.
2. Bebekler ağlar.
3. Köpekler konuşur.
4. Muzlar tatlıdır.

4

1. Armutlar mavidir.
2. Şapkalar başa giyilir.
3. Tavşanlar saati gösterir.
4. Filler büyüktür.

5

1. İnsanlar iki ayaklıdır.
2. Portakallar siyahtır.
3. Kediler futbol oynar.
4. Kitapları okuruz.

6

1. Tavşanlar ağaçta yetişir.
2. Biberler yeşildir.
3. Portakallar markette satılır.
4. İnsanlar üç gözlüdür.

5'LİK SETLER

1

1. Babalar kanatlıdır.
2. Dondurma soğuktur.
3. Portakallar gitar çalar.
4. Arabalar benzinle çalışır.
5. Fareler çok büyüktür.

2

1. Havuçlar mavidir.
2. Kulaklarımızla görürüz.
3. Portakallar turuncudur.
4. Tavuklar yumurta yapar.
5. Bıçak keskindir.

3

1. Elmalar pembedir.
2. Karıncalar yavaştır.
3. Dondurma sıcaktır.
4. Kediler fare kovalar.
5. Bebekler tüylüdür.

4

1. Kuşlar kocamandır.
2. Motorsikletler havlar.
3. Bıçaklar yumuşaktır.
4. Bulutlar beyazdır.
5. Tavuklar yazı yazar.

5

1. Gemiler uçar.
2. Kareler yuvarlaktır.
3. Çorabı ayağımıza giyeriz.
4. Bisikletler süt içer.
5. İnsanlar iki kulaklıdır.

6

1. Uçaklar kanatlıdır.
2. Elmalar şarkı söyler.
3. Dağlar çok küçüktür.
4. Sandalyeler ayaklıdır.
5. Makaslar kağıt keser.

6'LIK SETLER

1

1. Muzlar dişlidir.
2. Köpekler gitar çalar.
3. Bacağımız parmaklıdır.
4. Mektupları pulla göndeririz.
5. Muzlar sarıdır.
6. Kurbağalar zıplar.

2

1. Oyuncak ayılar yumuşaktır.
2. Ördekler suda yaşar.
3. Çocuklar üç kolludur.
4. Evimiz şarkı söyler.
5. Ördekler beş ayaklıdır.
6. Kar soğuktur.

3

1. Saatler zamanı gösterir.
2. Ayran tatlıdır.
3. Kurbağalar uzun kulaklıdır.
4. Ağaçlar müzik çalar.
5. Toplar yuvarlaktır.
6. Balıklar suda yaşar.

4

1. Arılar sokar.
2. Koyunlar kuyrukludur.
3. İnekler uçar.
4. Köpek balığı kocamandır.
5. Bulutlar siyahtır.
6. Pamuk ağırdır.

5

1. Ağaçlar tüylüdür.
2. Marketler yiyecek satar.
3. Domates kırmızıdır.
4. Kediler çok büyüktür.
5. Tavşanlar uzun kulaklıdır.
6. Tavuklar okula gider.

6

1. Kirazlar mavidir.
2. Ağaçlar yapraklıdır.
3. Demir hafiftir.
4. Yılanlar zıplar.
5. Kekler tatlıdır.
6. Tekerlekler karedir.

Experiment 4 (Reading Span Test)

Deneme seti

1. Kışın en soğuk zamanına zemheri denir.
2. Çorum İç Anadolu Bölgesi'nde yer almaktadır.
3. Almanya'da on iki milyon türk yaşamaktadır.

2'LİK SETLER

1

1. Senin kardeşinin çocuğu yiğenindir.
2. Trabzon mısırı ile ün salmıştır.

2

1. Haritada Türkiye Fransa'dan daha fazla yer kaplar.
2. 30 Eylül'de doğanlar akrep burcu olurlar.

3

1. Zorunlu eğitim ülkemizde 8 yıldır.
2. Uzağı iyi göremeyen hipermetrop gözlerdir.

4

1. Bir yumurta 80 kalori barındırır.
2. İnsan susuzluğa haftalarca dayanabilir.

5

1. Bir insanda 46 çift kromozom bulunur.
2. Türk Hukuk Kurumu THK ile kısaltılır.

6

1. Osmanlı İmparatorluğu 1299 yılında kurulmuştur.
2. Yapraklar ilkbaharda sararır.

3'LÜK SETLER

1

1. Salon sporlarından biri de bowlingdir.
2. Sebzeler bol miktarda B vitamini ihtiva eder.
3. Osmanlı Devleti dünyadaki en uzun süren imparatorluktur.

2

1. Boza içeceği Arap kökenlidir.
2. Bir bardak şekersiz çay sıfır kaloridir.
3. Sol ele söz yüzüğü takılır.

3

1. Lodos güneybatıdan esen rüzgara denir.
2. İskambil kağıdı ile briç oyunu oynanabilir.
3. Haritada Bulgaristan Yunanistan'dan daha fazla yer kaplar.

4

1. 1920 yılında cumhuriyet ilan edilmiştir.
2. Trampet nefesli bir çalgı türüdür.
3. Kıvırcık saçlı olmak kalıtımsaldır.

5

1. Uranüs güneşten en uzak olan gezegendir.
2. Rüştiye lise dereceli eğitim kurumuna denir.
3. Mozart Viyana'da doğmuştur.

6

1. Bir yıl üç yüz altmış beş gündür.
2. Çarparken çıkarmayı, bölerken toplamayı kullanırız.
3. Seksen tane şehir ülkemizde bulunmaktadır.

4'LÜK SETLER

1

1. 30 adet taşla tavla oyunu oynanabilir.
2. Etkisiz elemanı sıfır olan işlem toplamadır.
3. Türkiye'nin üçüncü cumhurbaşkanı Cemal Gürsel'dir.
4. İyot tiroit bezinin çalışması için gereklidir.

2

1. Baklagil türlerinden biri de mercimektir.
2. Dama ve satranç aynı sayıda taşla oynanmaktadır.
3. İzmir Muğla'dan yüzölçüm bakımından daha küçüktür.
4. Bir araba için hız sınırı otoyolda 90 km'dir.

3

1. Salep bir Türk içeceğidir.
2. Epik şiir kahramanlıklardan bahseder.
3. İç Anadolu Bölgesi Türkiye'nin en geniş bölgesidir.
4. 30 Mart'ta doğanlar kova burcu olurlar.

4

1. Tavşanlar ot yiyerek yaşar.
2. Sarı ve kırmızı birlikte karışırsa yeşil olur.
3. Yılan ve timsah sürüngendir.
4. % 74 oranında su çiğ yumurtada bulunur.

5

1. Kuş türlerinden biri de devekuşudur.
2. Teyzenin çocuğu senin yiğenindir.
3. 35 kalorilik enerji havuçta vardır.
4. Telli çalgılara örnek olarak akordiyon verilebilir.

6

1. Dünyanın yüzölçümü en büyük olan ülkesi Amerika'dır.
2. Ev telefonları elektrikle çalışır.
3. Azerbaycan Türkiye'ye komşudur.
4. Ay dünyanın üçte biri büyüklüğündedir.

5'LİK SETLER

1

1. Tuzlu su daha kısa sürede kaynamaktadır.
2. Bolu'nun yüzölçümü Sivas'ın yüzölçümünden büyüktür.
3. Oyun kartları 52 adet karttan oluşmaktadır.
4. Bir gözleri açık uyuyan hayvan yunuslardır.
5. Yirmi dört tane diş çocuklarda bulunmaktadır.

2

1. İstanbul 1453'de fethedildi.
2. İskambil kağıtlarındaki kupa ve sinek kırmızıdır.
3. C vitamini domateste bulunur.
4. Bir ünlem cümlesine örnek olarak aman tanrım verilebilir.
5. Jupiter güneşe en yakın gezegendir.

3

1. Rafting akarsuda yapılan bir spordur.
2. 15'şer adet siyah ve beyaz taş satrançta bulunur.
3. Dünya'nın en uzun insanı iki metre doksan cm boyundadır.
4. Fıstık fındıktan daha yağlı bir kuruyemiştir.
5. Roma rakamında C harfi ile 100 sayısı gösterilir.

4

1. Bir kilometre bir milden daha uzundur.
2. Tatlı su balıklarından biri de alabalıktır.
3. Bir yıl elli dört hafta sürmektedir.
4. Altı kişilik iki takımla voleybol oynanabilir.
5. Haritada Rusya Çin'e göre daha fazla yer kaplar.

5

1. İsim tamlamasına örnek olarak balın peteği verilebilir.
2. Futbol on iki kişilik iki takımla oynanır.
3. Sigara sağlığa yararlıdır.
4. Doğu Anadolu Bölgesi Malatya'yı da içermektedir.
5. Miyop gözler yakını iyi göremez.

6

1. Kemençe telli bir çalgı türüdür.
2. Mor doğada nadir bulunan renklerdendir.
3. Suriye'nin yönetim şekli cumhuriyettir.
4. Bir karınca kendi ağırlığının 20 katını taşıyabilir.
5. Çorum leblebi ile ünlüdür.

6'LIK SETLER

1

1. Gökkuşağının ortasında bulunan renk yeşildir.
2. 30 gün çeken aylardan biri de Mayıs ayıdır.
3. Bir şişe maden suyu bir kaloridir.
4. Osmanlı Devleti'nin para birimi akçedir.
5. Türkiye'nin en uzun akarsuyu Kızılırmak'tır.
6. Güneş dünyamızdan daha küçüktür.

2

1. Kediler sadece siyah beyaz görebilirler.
2. Elektrik akımı ölçüm birimi volttur.
3. Otizmde zekada gerilik yoktur.
4. Patates asit oranı yüksek bir sebzedir.
5. Bu yıl cumhuriyetin 84. yılını kutluyoruz.
6. Dünya'nın en yoğun nüfuslu ülkesi Hindistan'dır.

3

1. Poyraz sıcak bir rüzgar türüdür.
2. Antalya'nın nüfusu İstanbul'unkinden daha fazladır.
3. A vitamini göz sağlığı için gereklidir.
4. Türkiye Avrupa Konseyi'ne üye olmuştur.
5. Yeşil ve siyah renkler karışırsa kahverengi olur.
6. Toplam 184 ülke dünyada bulunmaktadır.

4

1. Elma asit oranı yüksek olan meyvelerdendir.
2. Kanın pıhtılaşması için kalsiyum gereklidir.
3. 8 kalorilik enerji salatalıkta vardır.
4. Mustafa Kemal Selanik'te doğmuştur.
5. Zebraların siyah üstüne beyaz çizgili derileri vardır.
6. Bursa'nın nüfusu Sivas'ın nüfusundan eksiktir.

5

1. Ankara'nın yüzölçümü Konya'nınkinden büyüktür.
2. Sıfat tamlamasına örnek olarak kapının kolu verilebilir.
3. Kılıçla yapılan sporlardan biri de eskrimdir.
4. Çiçekler kış mevsiminde açar.
5. Pirinç bir tahıl türüdür.
6. 26 tane harf İngilizlerde bulunmaktadır.

6

1. Mimar Sinan Türk soyundan gelmektedir.
2. Peynirde D vitamini bulunur.
3. Beş kişilik iki takımla basketbol oynanabilir.
4. Güneş sisteminde dokuz tane gezegen bulunmaktadır.
5. Malatya kayısı ile ünlüdür.
6. Ege Bölgesi Balıkesir'i de bulundurur.

APPENDIX B DEBRIEFING AND CONSENTS FOR ALL EXPERIMENTS

GÖNÜLLÜ KATILIM FORMU

Bu çalışma, ODTÜ Enformatik Enstitüsü Bilişsel Bilimler Bölümü'nde master yapmakta olan Gülten ÜNAL tarafından yürütülmekte olan bir tez çalışmasıdır. Çalışmanın amacı, çocuklarda ve karşılaştırmalı olarak yetişkinlerde ileriye doğru bozucu etkinin ve bunun beynin yönetici fonksiyonları ile olan ilişkilerinin incelenmesidir. İleriye doğru bozucu etki, bir konuyu öğrenirken o konudan hemen önce öğrenilmiş başka bir konunun, esas konuyu öğrenmeyi zorlaştırmasına denir.

Bu çalışmayla beraber, katılımcılar kendi bilişsel gelişimlerini inceleme fırsatı bulabileceklerdir. Öğretmenler, bu deneylerin sonuçlarına bakarak öğrencilerinin bilişsel gelişimleri hakkında faydalı bilgilere sahip olabilecektir. Buna ek olarak, öğrencinin ailesi de öğrencideki bilişsel gelişmeyi inceleyip, çocuğuna destek olabilir.

Çalışmada katılımcılara 4 tane davranışsal hafıza deneyi yapılcaktır. Bunlardan 3 tanesinde, katılımcıdan beklenen şey sadece hatırladığı kelimeleri söylemek olacaktır. Verilen cevaplar bir ses kayıt cihazı yardımıyla kaydedilecektir. Diğer 4. deneyde ise katılımcıdan bazı özel oyun kartlarını kullanarak, tahminlerde bulunması istenmektedir. İlk üç test ortalama 5'er dakika sürmekte olup, en son yapılacak olan deneyin süresi ise, katılımcının performansına göre 10-20 dakika arasında değişmektedir.

Bu çalışmaya katılmak tamamen gönlüllük çerçevesinde olmalıdır. Çalışmada katılımcıya rahatsızlık verebilecek unsurlar bulunmamaktadır. Fakat, çalışma sırasında deneylerden ya da diğer herhangi bir sorundan ötürü kendinizi rahatsız hisserdeniz, çalışmayı yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durum sözkonusu olduğunda, deneyi uygulayan kişiye bunu söylemeniz yeterli olacaktır. Ayrıca bu çalışmaya katılmakta katılımcılar için herhangi bir riski bulunmamaktadır.

Bu çalışma sırasında elde edilecek olan bilgiler tamamiyle gizli tutulacak ve sadece bu araştırmayi yürütenler tarafından kullanılacaktır.

Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için Bilişsel Bilimler Bölümü öğretim üyesi Dr. Annette Hohenberger (Tel: E-posta:) ve Bilişsel Bilimler Bölümü master öğrencisi Gülten ÜNAL (Tel:; E-posta:) ile iletişim kurabilirsiniz.

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

VELİ ONAY FORMU

Sayın Veli,

Orta Doğu Teknik Üniversitesi Enformatik Enstitüsü Bilişsel Bilimler Bölümü'nde yüksek lisans öğrencisi olarak çalışmaktayım.Master tezim kapsamında "İleriye doğru bozucu etkiden kurtulma ve bunun yönetici fonksiyonlarla olan ilişkileri: Türk çocuklarında gelişimsel bir çalışma" başlıklı bir tez çalışması yürütmekteyim. Bu çalışmanın amacı 7-12 yaş grubu çocuklarda ileriye doğru bozucu etkinin yaşla doğru orantılı olarak azalmasını incelemektir. (İleriye doğru bozucu etki, bir konuyu öğrenirken, o konudan hemen önce öğrenilmiş başka bir konunun, esas konuyu öğrenmeyi zorlaştırmasına denir.) Bu amacı gerçekleştirebilmek için çocuklarınızla bazı "davranışsal hafıza deneyleri" yapmaya ihtiyaç duymaktayım.

Bu çalışmada çocuklara 4 tane davranışsal hafıza deneyi yapılcaktır. Bunlardan 3 tanesinde, çocuklardan beklenen şey sadece hatırladığı kelimeleri söylemek olacaktır. Verilen cevaplar bir ses kayıt cihazı yardımıyla kaydedilecektir. Diğer 4. deneyde ise bazı özel oyun kartlarını kullanarak, çocuklardan tahminlerde bulunması istenmektedir. İlk üç test ortalama 5'er dakika sürmekte olup, en son yapılacak olan deneyin süresi ise, çocuğun performansına göre 10-20 dakika arasında değişmektedir.

Çalışmaya katılım tamamiyle gönüllülük çerçevesindedir. Hem sizin onayınız hem de bu çalışmaya katılması için çocuğunuzun gönüllü olması bir ön şarttır. Katılmasına izin verdiğiniz takdirde deneyleri okulda ders saatinde gerçekleştireceğiz. Çocuğunuzun katılacağı deneylerin onun psikolojik gelişimine olumsuz etkisi olmayacağından emin olabilirsiniz. Çocuğunuz bu deneylerdeki cevapları kesinlikle gizli tutulacak ve bu cevaplar sadece bilimsel araştırma amacıyla kullanılacaktır. Katılım sonunda, herhangi bir maddi yarar sağlanmamakatdır. Bu formu imzaladıktan sonra çocuğunuz katılımcılıktan ayrılma hakkına sahiptir. Çalışma sırasında da çocuğunuz herhangi bir sebepten ötürü çalışmayı yarıda bırakmakta serbesttir.

Çocuğunuzun deneylere katılarak bize sağlayacağı bilgiler çocukların bilişsel gelişimlerini incelemek adına önemli katkılarda bulunacaktır. Araştırmayla ilgili sorularınızı aşağıdaki e-posta adresini veya telefon numarasını kullanarak bize yöneltebilirsiniz.

Saygılarımızla,

İmza | İmza

Dr. Annette Hohenberger | Gülten Ünal

Lütfen bu araştırmaya çocuğunuzun katılması konusundaki tercihinizi aşağıdaki seçeneklerden size en uygun geleni daire içine alarak ve *imzanızı atarak belirtiniz ve bu formu çocuğunuzla okula geri gönderiniz.*

Yukarıda açıklamasını okuduğum çalışmaya, oğlum/kızım _____________________’nin katılımına izin veriyorum/ izin vermiyorum. Ebeveynin:

Adı, soyadı: __________________ İmzası: ____________ Tarih: ______________

KATILIM SONRASI BİLGİ FORMU

Bu çalışma, ODTÜ Enformatik Enstitüsü Bilişsel Bilimler bölümünde master yapmakta olan Gülten ÜNAL tarafından yürütülmekte olan bir tez çalışmasıdır. Çalışmanın amacı, daha önce de belirtildiği gibi, çocuklarda ve karşılaştırmalı olarak yetişkinlerde ileriye doğru bozucu etkinin ve bunun beynin yönetici fonksiyonları ile olan ilişkilerinin incelenmesidir. İleriye doğru bozucu etki, bir konuyu öğrenirken o konudan hemen önce öğrenilmiş başka bir konunun, esas konuyu öğrenmeyi zorlaştırmasına denir.

Çalışmada 4 farklı davranışsal hafıza deneyi bulunmaktadır. Ana deney olarak kategorisel kelime hatırlama deneyi yapılcaktır. Bu deneyde birbiri ardınca gelen 3 farklı kategori bulunmaktadır. Herbir kategoride öğrenilen kelimeler, bir sonraki kategorinin öğrenilmesiyle beraber unutulmaya başlanmaktadır. Bu durumda yeni kategorideki kelimeleri aklınızda tutmaya çalışırken, bir önceki kategoride öğrendiğiniz kelimeler yeni kelimeleri aklınızda tutmanızı zorlaştıracaktır. Ayrıca, herbir yeni kategorinin ilk kelimesinin, kategori değişimi yaşandığı için, hatırlanması kolaylaşacaktır. Bunun dışındaki diğer 3 deneyde de, beynin yönetici fonksiyonları yine hafıza testleriyle ölçülmüştür. Bu ek deneylerin yapılmasındaki amaç, beynin yönetici fonskiyonları ile ilerye doğru bozucu etkiden kurtulmanın arasındaki ilişkileri gözlemlemektir.

Bu çalışmadan elde edilen bilgiler sadece bilimsel araştırma ve yazılarda kullanılacaktır. Çalışmanın sonuçlarını öğrenmek ya da bu araştırma hakkında daha fazla bilgi almak için aşağıdaki isimlere başvurabilirsiniz. Bu araştırmaya katıldığınız için tekrar çok teşekkür ederiz.

APPENDIX C QUESTIONNAIRES FOR TEACHERS

VELİ ARAŞTIRMA ANKETİ

S1 Annenin eğitim durumu

☐ hiç bir okulu bitirmemiş ☐ ilkokul ☐ ortaokul ☐ lise
☐ meslek lisesi, hangi alanda?________________________
☐ Üniversite, derece, alan?____________________________
☐ diğer ____

S2 Anne şu anda çalışıyor mu?

☐ hayır ☐ evet
Evetse, Ne iş yapıyor?________________________________

S3 Babanın eğitim durumu

☐ hiç bir okulu bitirmemiş ☐ ilkokul ☐ ortaokul ☐ lise
☐ meslek lisesi, hangi alanda?________________________
☐ Üniversite, derece, alan?____________________________
☐ diğer ____

S4 Baba şu anda çalışıyor mu?
☐ hayır ☐ evet Evetse, Ne iş
yapıyor?____

S5 Aile durumu:
☐ evli ☐ ayrılmış/boşanmışlar

S6 Bu çalışmaya katılan çocuğun kaç kardeşi var?
☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 veya daha fazla
Bu çalışmaya katılan çocuğunuz kaçıncı
çocuğunuz?_________________________________

ÖĞRENCİ BAŞARI ANKETİ

S1 Bu öğrencinin ne kadar zeki olduğunu düşünüyorsunuz?

☐ 1 (çok düşük) ☐ 2 (düşük) ☐ 3 (orta) ☐ 4 (iyi) ☐ 5 (çok iyi)

S2 Sizce öğrencinin hafızası ne kadar iyidir?

☐ 1 (çok düşük) ☐ 2 (düşük) ☐ 3 (orta) ☐ 4 (iyi) ☐ 5 (çok iyi)

S3 Sizce öğrencinin anlatılanları anlama kapasitesi ne kadar iyidir?

☐ 1 (çok düşük) ☐ 2 (düşük) ☐ 3 (orta) ☐ 4 (iyi) ☐ 5 (çok iyi)

S4 Sizce öğrencinin kendi ana dilini kullanabilme kabiliyeti hangi seviyededir?

☐ 1 (çok düşük) ☐ 2 (düşük) ☐ 3 (orta) ☐ 4 (iyi) ☐ 5 (çok iyi)

S5 Sorulanlar dışında öğrencinin başka herhangi bir problemi var mı?

☐ Hayır

☐ Evet Evetse, Bu problem nelerdir?____

____

APPENDIX D MEAN TABLES FOR THE CFR-TASK

Table 69 Table of means for all categories & items, gradewise

			Mean	**S. D.**	**N**
Grade 1	*Category #1*	item 1	,35	,489	20
		item 2	,45	,510	20
		item 3	,60	,503	20
		item 4	,20	,410	20
	Category #2	item 1	,40	,503	20
		item 2	,35	,489	20
		item 3	,25	,444	20
		item 4	,30	,470	20
	Category #3	item 1	,60	,503	20
		item 2	,40	,503	20
		item 3	,50	,513	20
		item 4	,25	,444	20
Grade 2	*Category #1*	item 1	,38	,495	24
		item 2	,71	,464	24
		item 3	,54	,509	24
		item 4	,29	,464	24
	Category #2	item 1	,54	,509	24
		item 2	,25	,442	24
		item 3	,25	,442	24
		item 4	,38	,495	24
	Category #3	item 1	,50	,511	24
		item 2	,50	,511	24
		item 3	,63	,495	24
		item 4	,33	,482	24
Grade 3	*Category #1*	item 1	,63	,500	16
		item 2	,69	,479	16
		item 3	,50	,516	16
		item 4	,25	,447	16
	Category #2	item 1	,38	,500	16
		item 2	,38	,500	16
		item 3	,38	,500	16
		item 4	,56	,512	16
	Category #3	item 1	,38	,500	16
		item 2	,50	,516	16
		item 3	,63	,500	16
		item 4	,56	,512	16
Grade 4	*Category #1*	item 1	,86	,351	22
		item 2	,59	,503	22
		item 3	,32	,477	22

Table 69 (continued)

			Mean	S. D.	N
		item 4	,59	,503	22
	Category #2	item 1	,50	,512	22
		item 2	,41	,503	22
		item 3	,32	,477	22
		item 4	,27	,456	22
	Category #3	item 1	,55	,510	22
		item 2	,55	,510	22
		item 3	,55	,510	22
		item 4	,36	,492	22
Grade 5	*Category #1*	item 1	,84	,375	19
		item 2	,63	,496	19
		item 3	,58	,507	19
		item 4	,37	,496	19
	Category #2	item 1	,63	,496	19
		item 2	,42	,507	19
		item 3	,21	,419	19
		item 4	,37	,496	19
	Category #3	item 1	,37	,496	19
		item 2	,53	,513	19
		item 3	,37	,496	19
		item 4	,21	,419	19

Table 70 Table of means for all categories & items, across gender (adults)

			Mean	**S. D.**	**N**
Females	*Category #1*	item 1	1,00	,000	11
		item 2	1,00	,000	11
		item 3	1,00	,000	11
		item 4	,55	,522	11
	Category #2	item 1	,82	,405	11
		item 2	,91	,302	11
		item 3	,64	,505	11
		item 4	,73	,467	11
	Category #3	item 1	,91	,302	11
		item 2	,91	,302	11
		item 3	,73	,467	11
		item 4	,64	,505	11
Males	*Category #1*	item 1	1,00	,000	9
		item 2	,78	,441	9
		item 3	,89	,333	9
		item 4	,67	,500	9
	Category #2	item 1	,78	,441	9
		item 2	1,00	,000	9
		item 3	,67	,500	9
		item 4	,33	,500	9
	Category #3	item 1	,44	,527	9
		item 2	,78	,441	9
		item 3	,44	,527	9
		item 4	,33	,500	9

Table 71 Table of means for all categories & items (5th graders and adults)

			Mean	**S. D.**	**N**
Grade 5	*Category #1*	item 1	,84	,375	19
		item 2	,63	,496	19
		item 3	,58	,507	19
		item 4	,37	,496	19
	Category #2	item 1	,63	,496	19
		item 2	,42	,507	19
		item 3	,21	,419	19
		item 4	,37	,496	19
	Category #3	item 1	,37	,496	19
		item 2	,53	,513	19
		item 3	,37	,496	19
		item 4	,21	,419	19
Adults	*Category #1*	item 1	1,00	,000	20
		item 2	,90	,308	20
		item 3	,95	,224	20
		item 4	,60	,503	20
	Category #2	item 1	,80	,410	20
		item 2	,95	,224	20
		item 3	,65	,489	20
		item 4	,55	,510	20
	Category #3	item 1	,70	,470	20
		item 2	,85	,366	20
		item 3	,60	,503	20
		item 4	,50	,513	20