Characteristics and Neuronal Correlates of Superior Memory Performance

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Philosophie an der Ludwig-Maximilians-Universität

München

vorgelegt von

Boris Nikolai Konrad

aus Bochum

München, Oktober 2013

Referentin: Prof. Dr. Kristina Hennig-Fast

Koreferent: Prof. Dr. Patrick Khader

Tag der mündlichen Prüfung: 24.01.2014

Acknowledgment

It was several years ago that I first talked to Martin Dresler about the possibility of studying superior memorizers in order to learn more about their skill and memory in general. I was still a physics and computer science student at the time and had no idea that large parts of what we discussed would eventually end up in my PhD thesis. I would hereby like to express my deepest gratitude to everyone who supported, assisted and motivated me during the last approximately four years that it took me since starting at the Max Planck Institute of Psychiatry to finish this thesis, and I also extend my gratitude to everyone who contributed to the studies.

In particular I would like to thank Michael Czisch for giving me the opportunity to undertake this work within his research group, Kristina Hennig-Fast for supervising my work and providing many helpful comments and ideas during the progress, Patrick Khader for co-supervising this thesis and Axel Steiger for assisting with the sleep-related parts of the study and having me in his research group in the last few months.

Martin Dresler first had the idea to study memory athletes, encouraged me to join him and supported me on each step to realizing the various studies that eventuated from that initial idea. Thank you, Martin!

I also would like to thank Lisa Genzel and Victor Spoormaker for their endless helpfulness and willingness to comment on all parts of my work, answering any and all questions that I had. This is also true for each and every person in the RGs Czisch and Steiger, including the technicians in both groups who provided a lot of help with equipment, measurements and sleep scoring. Thanks to all of you for your many helpful comments and encouraging words in group meetings, journal clubs and beyond. Thank you also to everyone who commented on prior versions of this thesis.

Hannah Nickl conducted the learning sessions and sleep parts of Study 3 and I would like to thank her for her support and her patience. Laurentius Halimsetiawan performed analysis of the sleep data in Study 1.

I would also like to thank all of the subjects who joined me and were patient and flexible when needed. Over 150 in total (counting all pilots and controls) participated; some of them even came all the way from Sweden and Norway despite full schedules as full-time doctors or lawyers! Thank you!

Besides work-related support I am also grateful for the support I received from many friends and of course my family, in particular my parents. Thanks to all of you for always being there when I needed someone to talk to.

Finally I like to thank the financial sponsors: The Friedrich-Ebert-Stiftung (FES) who funded me with a PhD student scholarship and the Max Planck Institute of Psychiatry, who covered all of the study-related costs.

Zusammenfassung

„Jedermann klagt über sein Gedächtnis, niemand über seinen Verstand."

(François de La Rochefoucauld, ca. 1664)

Viele Menschen fürchten sich vor Gedächtnisverlust im Alter und sind auch im Alltag mit ihrem Gedächtnis nicht zufrieden. Im Gegensatz dazu gibt es einzelne Personen die mit außergewöhnlich guten Gedächtnisleistungen imponieren. Sie zeigen ihre Fähigkeiten in Fernsehshows oder bei Gedächtnismeisterschaften und werden dafür oft bewundert. Schon im Mittelalter sind Gedächtniskünstler aufgetreten (Yates, 1966) und Wissenschaftlern (Valentine & Wilding, 1997) aufgefallen. Ihre Leistungen basieren zumeist auf Mnemotechnik genannten Gedächtnistechniken (Maguire, Valentine, Wilding, & Kapur, 2003), die schon seit der Antike bekannt sind (Hrees, 1986). Dennoch ist der Umfang der wissenschaftlichen Forschung über Gedächtnisverbesserung und Mnemotechniken im Vergleich zu Studien über Gedächtniserkrankungen noch stark unterrepräsentiert (Worthen & Hunt, 2011).

Im Jahr 1990 wurde die erste Gedächtnisweltmeisterschaft ausgetragen. Seitdem hat sich der Gedächtnissport etabliert und es finden inzwischen weltweit entsprechende Meisterschaften statt (Wilding & Valentine, 1994). Erstmalig wurden vor rund zehn Jahren einige Gedächtnissportler auch mit Hilfe von bildgebenden Verfahren (funktionelle Magnetresonanztomographie, fMRT) untersucht (Maguire, Valentine, et al., 2003). Seitdem haben sich Teilnehmerzahlen und Rekorde (etwa das Merken der Reihenfolge eines gemischten Kartenspiels in 21,19 Sekunden) im Gedächtnissport rasant entwickelt (Konrad, 2013). In der ersten Studie wurden insgesamt 28 Gedächtnissportler in einer Reihe von Teilstudien untersucht. Dabei konnte gezeigt werden, dass anders als noch von Maguire und Kollegen berichtet, die Gedächtnissportler überdurchschnittlich intelligent waren. Insbesondere fielen ihre Leistungen in einer Aufgabe zur Messung der Verarbeitungsgeschwindigkeit auf. Selbst im Vergleich zu einer Kontrollgruppe gleicher Intelligenz waren die Gedächtnissportler besser und zudem korrelierten ihre Bestleistungen in der schnellsten Disziplin im Gedächtnissport (Speed Cards) mit der

Verarbeitungsgeschwindigkeit. Mit Hilfe zweier Gedächtnisaufgaben (Directed Forgetting und False Memories) konnte gezeigt werden, dass die Gedächtnissportler nur dann außerordentliche Leistungen zeigen, wenn sie auf ihre Techniken zurückgreifen. -Sofern dies geschieht, können sie sich nicht nur mehr merken, sondern haben auch einen genaueren Zugriff auf Gedächtnisinhalte und sind weniger anfällig für Gedächtnisfehler. Im Rahmen einer fMRT-Untersuchung wurde die Wiedergabe von Folgen von Binärziffern untersucht, die entweder einige Tage zuvor oder direkt vor der Wiedergabe auswendig gelernt wurden. Hierbei zeigte sich, dass selbst bei gerade erst mit Hilfe der Mnemotechniken gelernten Ziffernfolgen keine Aktivierung von mit Arbeitsgedächtnis-Gedächtnisprozessen assoziierten fronalten Gehirnarealen feststellbar ist. Im Gegensatz stehen Folgen kurzer Ziffern, die mit dem Arbeitsgedächtnis gelernt wurden, sowie andere Studien, die zeigen, dass bei Verwendung von Gruppierungsstrategien im Arbeitsgedächtnis die frontale Aktivierung noch zunimmt (Bor, Duncan, Wiseman, & Owen, 2003). Diese Ergebnisse unterstützen das Modell des Langzeitarbeitsgedächtnisses (Ericsson & Kintsch, 1995), welches postuliert, dass Experten in ihrem Themengebiet direkt auf Netzwerke im Langzeitgedächtnis zurückgreifen können, wenn sie dazu passende neue Inhalte lernen. Dies geschieht im gleichen Tempo wie sonst nur beim Arbeitsgedächtnis. Dazu passt auch, dass bei Arbeitsgedächtnisaufgaben selbst keine Unterschiede zwischen Gedächtnissportlern und Kontrollgruppen zu finden waren.

Obwohl die Gedächtnissportler ausnahmslos alle angegeben haben, dass ihre außergewöhnlichen Leistungen auf den Mnemotechniken und intensivem Training basieren, kann allein auf diesen Daten nicht herausgefunden werden, welche etwaige Fähigkeiten möglicherweise nötig sind, um solche Leistungen zu erreichen. Deshalb wurden in der zweiten Studie normale Probanden in Gedächtnistechniken trainiert. Dabei standen die Routenmethode (Roediger, 1980) und das Major-System (Patton, 1986) im Vordergrund, welche sich beide in zahlreichen Studien bereits als wirksam erwiesen haben (Worthen & Hunt 2011). Allerdings gab es bisher nur wenige Studien, die Gedächtnistraining über einen längeren Zeitraum begleitet haben (Ericsson, Chase, & Faloon, 1980; Higbee, 1997; Kliegl, Smith, Heckhausen, & Baltes, 1987). Diese blieben zudem auf die Untersuchung von Einzelpersonen begrenzt. Die vorliegende Studie war so die erste, welche ein intensives Gedächtnistraining mit einer Gruppe von Probanden begleitet hat. 20 Probanden haben an einem zweitägigen Kurs zu Gedächtnistechniken teilgenommen und anschließend über sechs bis acht Wochen mittels einer Online-Plattform von zu Hause aus trainiert. Insgesamt 13 der Probanden erfüllten das Trainingssoll von mindestens 20 Stunden. Dabei zeigten sich deutliche Trainingsverbesserungen in den Gedächtnisaufgaben. So konnten die Teilnehmer in einer Zahlenmerkaufgabe ihre Leistung mehr als verdoppeln und auch im Wörtermerken deutliche Verbesserungen erzielen. Dabei waren die erzielten Verbesserungen unabhängig von der Intelligenz oder der Leistungsfähigkeit zu Beginn. Darüber hinaus verbesserten sich die Trainingsprobanden auch in der Transferaufgabe zur Verarbeitungsgeschwindigkeit signifikant, im Gegensatz zur Kontrollgruppe, die nicht trainiert hat. Interessanterweise stieg die Selbsteinschätzung der eigenen Gedächtnisleistung trotz großen messbaren Erfolgs nur gering. Dies ist ein Indiz dafür, warum sich in Umfragen zeigt, dass viele trotz Kenntnis von Gedächtnismethoden diese nicht anwenden (Soler, María JoseRulz, 1996).

In der dritten Studie wurden Probanden nur relativ kurz (eine Stunde) in der Routenmethode instruiert und sollten diese anschließend benutzen, um sich 50 Begriffe in der richtigen Reihenfolge einzuprägen. Danach hielten sie für eine Stunde Mittagsschlaf oder blieben wach. Aus der Forschung ist bekannt, dass die Gedächtnisleistung von Schlaf profitiert (Rasch & Born, 2013) und dies konnte auch für mittels der Routenmethode gelernte Begriffe gezeigt werden. Wenn im Schlaf zusätzliche Töne eingespielt wurden, die mit den gelernten Begriffen assoziiert worden waren (Rudoy, Voss, Westerberg, & Paller, 2009), konnten die Probanden diese Begriffe hinterher besser wiedergeben. Allerdings ging dies zu Lasten der nicht unterstützten Begriffe, was dafür spricht, dass die im Schlaf stattfindenden Prozesse der Gedächtniskonsolidierung durch äußeren Einfluss einseitig beeinflusst werden, wie schon durch eine Tierstudie gezeigt (Bendor & Wilson, 2012).

Zusammenfassend lässt sich aufgrund dieser Arbeit festhalten, dass außergewöhnlich gute Gedächtnisleistungen mit Hilfe von Gedächtnistechniken auch für normale Probanden in wenigen Wochen erreichbar sind (Studie 2), dies mit einer Verbesserung auch der Verarbeitungsgeschwindigkeit einhergeht (Studie 2), die auch bei Gedächtnissportlern äußert gut ist und zudem mit der Gedächtnisleistung korreliert (Studie 1). Durch das Gedächtnistraining wird das Arbeitsgedächtnis nicht beeinflusst (Studie 1 und 2), dafür aber werden Langzeitgedächtnisstrukturen so trainiert, dass sie im gleichen Tempo wie sonst nur das Arbeitsgedächtnis beschrieben werden können (Studie 1), was sich auch mittels funktioneller Bildgebung zeigt (Studie 1). Einen Einfluss auf die Gedächtniskonsolidierung im Schlaf hat dies allerdings nicht (Studie 1 und Studie 3). Die Routenmethode kann auch nach kurzer Instruktion sofort gewinnbringend eingesetzt werden (Studie 3) und kann in verschiedenen Gedächtnisaufgaben und unter unterschiedlichen Modalitäten eingesetzt werden (Studie 1).

Abstract

„Everyone complains of his memory, and no one complains of his judgment."

(François de La Rochefoucauld, ca. 1664)

Many are afraid of memory loss in ageing and are additionally unsatisfied with their memory for everyday tasks. In contrast there are a few individuals who are capable of superior memory performances. They appear on TV and compete in memory competitions and garner the admiration of others. Memory performers have existed since the middle ages (Yates, 1966) and have been studied by scientists (Valentine & Wilding, 1997) before. Their performances are usually underpinned by mnemonic techniques (Maguire, Valentine, et al., 2003) which have been in use since ancient times (Hrees, 1986). However, the amount of scientific studies on memory improvement and mnemonic techniques is very limited compared to studies on dementia and memory loss (Worthen & Hunt, 2011).

The first World Memory Championships were run in 1990. Since then memory sports have become established and nowadays many competitions are run around the world (Wilding & Valentine, 1994). In 2003 a study on memory athletes employing neuroimaging methods (functional magnetic resonance imaging, fMRI) was published for the first time (Maguire, Valentine, et al., 2003). During the last ten years the number of competitors and records (pertaining to, for example, memorizing the order of a shuffled deck of playing cards within 21.19 seconds) in memory sports has risen rapidly (Konrad, 2013). Therefore in the first study the abilities of 28 memory athletes were assessed across a range of different tasks. It could be seen that, in contrast to the findings of Maguire and colleagues, memory athletes also possessed above average intelligence. In particular, their performance in a processing speed task was extraordinary even when compared to intelligence-matched controls, where performances in the fastest memory sports event (Speed Cards) correlated with processing speed.

Using two memory tasks (directed forgetting and false memories) it was shown that memory athletes are only able to show superior performances when they apply their methods. If they

do so, then they show more than just superior memory capacity; they also exhibit higher memory accuracy and less vulnerability for memory errors. Using fMRI the recall of binary digits was investigated, where digits had either been learnt days before, or immediately before retrieval. This investigation revealed that, even for recently learned binary digits, there was no activation in the frontal brain regions that are usually associated with working memory when mnemonics were applied; however the pattern of brain activation mirrored that seen in the retrieval of older binaries. Frontal working memory areas were activated only for short sequences of binaries. This finding is in contrast to studies showing that applying chunking methods to store more data in working memory even increases frontal activity (Bor et al., 2003). Therefore these findings support the long-term working memory theory (Ericsson & Kintsch, 1995), which postulates that experts do build networks in long-term memory in their area of expertise, which they can then utilize to learn related information at a pace at which, normally, only working memory can be accessed. Congruent with this finding was the fact that memory athletes and controls did not differ in performance in the working memory tasks.

Even though all memory athletes reported that their skill is based on using, and intensely training in, mnemonic techniques, based on our findings for the athletes alone no judgment can be made as to which other abilities or characteristics might be necessary in order to achieve such performances. Therefore in Study 2 regular subjects were trained in mnemonic techniques. Training focused on the method of loci (Roediger, 1980) and also the phonetic mnemonic (Patton, 1986), which have both been shown to be effective before. However, thus far few studies have looked into prolonged memory training (Ericsson et al., 1980; Higbee, 1997; Kliegl et al., 1987). These studies were all concerned with a few individuals only. Therefore the present study is the first to investigate intensive mnemonic training with a group of regular subjects. 20 subjects joined a two day course in mnemonic techniques followed by six to eight weeks of at-home training using an online platform. 13 subjects fulfilled the training criterion of a minimum of 20 hours. Strong improvements in the memory tasks were found, where participants more than doubled their performance in memorizing digits and also showed strong improvements in a word memorization task. These improvements were independent of intelligence and pre-training memory abilities.

In addition the training subjects also improved their performance in a transfer task, as it pertains to processing speed, significantly more so than did controls, who did not train. Interestingly the self-assessment of their memory performance only mildly increased despite the marked memory enhancement which they achieved. This might be one reason why people often don't apply mnemonics even when they are cognizant of them (Soler, María JoseRulz, 1996).

In study three subjects were briefly trained (one hour) in the method of loci. They applied it such that they could memorize 50 words in order, followed by a nap or staying awake. It is known that memory benefits from sleep (Rasch & Born, 2013) and this was also the case for words memorized using the method of loci. When sounds were played during sleep that had been associated with words (Rudoy et al., 2009), subjects could retrieve more of the cued words. However, this was at expense of the uncued items, indicating that memory consolidation processes during sleep are only biased by external factors, as suggested by an animal study (Bendor & Wilson, 2012).

In summary this thesis shows that superior memory performance can be achieved using mnemonic techniques by regular subjects within a few weeks (study 2), and further that this improvement is aligned with a transfer to processing speed (study 2), which is also superior in memory athletes, and is even correlated with memory speed (study 1). Mnemonic training did not influence working memory (study 1 and 2), but long-term memory structures develop, which can be assessed at a pace at which, normally, only working memory can be assessed at (study 1). This could be further supported by fMRI findings (study 1). However this does not influence memory consolidation in sleep (study 1 and 3). The method of loci can be utilized after brief instruction (study 3) and can be used for various memory tasks, and in different modalities (study 1).

Content

1. General Introduction

What we know and what we remember is essential for the formation of our personality and how we deal with the world. Humans fear memory loss, and not just in the context of ageing, and seek ways to improve their memory. But forgetting is part of the brain's way of organizing information. In particular, abstract details and contextual information are hard to remember. However some individuals demonstrate far superior memory abilities to the average person. They amaze the public with their performances and have garnered attention from cognitive scientists ever since the inception of that field of knowledge. Due to the limited number of individuals with these superior memorizing capabilities, the number of studies on them remains rather small, in particular compared to the vast amount of literature on Alzheimer's' disease and other forms of dementia. In particular group studies have rarely been conducted, and it has been suggested that it might be hard to follow up on existing studies with larger cohorts:

> "… it is unlikely that future studies would be able to recruit enough world-class memory performers to provide tests with much greater statistical power." (Ericsson, 2003)

With the introduction of memorizing as a competitive sport in the late 20th century, those individuals who demonstrate memory capabilities far above the norm got a platform via which they could meet and compete for prizes, and improve records for various memory tasks year upon year. Most of these individuals base their performance on the use of mnemonic strategies and training rather than innate ability (Konrad & Dresler, 2007).

Besides long term memory, training in mnemonic techniques also improves performance in working memory tasks (Carretti, Borella, & De Beni, 2007), an effect which perseverates over a span of at least five years (Gross & Rebok, 2011). Worthen and Hunt (2011) recently published a broad overview on the topic, demonstrating that mnemonic techniques have gotten less attention since 1980, and further appealing for new research on the relationship between basic memory research and mnemonics, for which they suggest the new term "mnemonology".

The aim of this thesis is to investigate the phenomenon of superior memory abilities. Three studies have been conducted for this purpose. The first study compares superior memorizers, as identified by their success in memory competitions, to matched control subjects via behavioral testing as well as sleep assessments and brain imaging. For the second study subjects were extensively instructed in mnemonic techniques and were required to practice them for at least six weeks afterwards. Behavioral testing, as well as brain imaging, was conducted before and after initial instruction, as well as after the training phase, and was compared to an untrained wait list control group. For the third study, subjects received only a brief introduction to a specific mnemonic technique known as method of loci. Here, subjects had to memorize lists of words using the method of loci and subsequently had to recall the remembered items in the brain scanner before and after either an afternoon nap or a period of wake.

1.1. Individuals with superior memory abilities

Memory refers to the process of information encoding, storage and retrieval. Because we can process various kinds of information, there are various kinds of memory. Furthermore, as there are various kinds of memory, there are various kinds of individuals with superior memory performance. In their book "Superior Memory", John Wilding and Elizabeth Valentine use three criteria to define superior memory ability: "(1) rapid acquisition of material or (2) acquisition of an unusually large quantity of material in a measured time, and (3) long-term retention of an unusually large quantity of material acquired under controlled conditions." (Valentine & Wilding, 1997). This definition will also be used within the present thesis.

1.1.1. Single Case Studies

In their book, Wilding and Valentine review the scientific literature on individuals known for their exceptionally good memory. Early reports are brief and anecdotal, and date back to the first century BC. The first reports which give actual figures describing superior memory performers appeared in the 18th century. In 1894 the French psychologist Alfred Binet was the first to study some of these individuals, using defined tests with differential modalities of presentation. In the following decades, several more individuals with superior memory performance have been tested around the world, and by various psychologists. As different kinds of tasks have been used, and usually in the absence of reportage on the specific material employed, it is hard to compare these individuals. The fairest comparison which can be made is via digit memorization tasks, which are the most culturally-neutral paradigm available, which furthermore can be summarized in terms of the speed of memorization. However, even in this task the form of presentation, and the memorization time, can vary markedly, which must be taken into consideration when comparing scores. Wilding and Valentine list the performance in digit memorization tasks of 16 different individuals identified through their literature review. The best-performing individual of those was the Japanese stage performer Ishihara (Susukita, 1933), who achieved the longest string of numbers memorized under controlled settings (2502 digits at a memorization speed of 5.3s per item with 89% correctly recalled as well as 2400 digits at a speed of 6.2s per item with 98% correctly recalled).

One of the most famous subjects in this category is Solomon Veniaminovich Shereshevsky, often simply referenced as “S.”. S. was a Russian journalist with an unusual skill for memory, who was intensively studied by Alexander Luria during the late 1920s and 1930s. The results were first published in 1968 and remain one of the most famous accounts available on superior memory (Luria, 1968). Luria diagnosed S. with a strong case of synesthesia. When memorizing digits, he made use of the spontaneous associations he had with numbers. However, and in contrast to other superior memorizers, S. struggled to control these associations, which sometimes disturbed other cognitive processes. He also had problems forgetting unnecessary information. Luria also reports that S. was able to recall strings of digits and other items many years later, upon being retested without prior warning. S. also made use of locations, visualizing to-be-remembered items superimposed upon them. Despite the strong similarity between S.´s way to remember and ancient memory techniques, Luria did not elaborate on this connection and instead concluded that S. might store visual representations of everything he saw. It was subsequently shown that Luria's own data contradict this assumption and that despite the unusual struggles S. had with his memory and his lack of introduction into mnemonic techniques, the way he actually stored information is comparable to the techniques employed by modern day memorizers (Valentine & Wilding, 1997). In general, these case reports differ with respect to the generalizability of the memory skills of their subjects since reports on exceptional performances in everyday memory tasks by the studies subjects often remain anecdotal. In particular, reports on some of the memory artists who performed for the public show that they often only displayed superior memory ability for specific material.

Another case of an individual with a very specific memory talent pertains to a woman named Elizabeth (Stromeyer & Psotka, 1970). She claimed to possess a perfect eidetic memory, which was tested by random dot stereograms. This task had been developed to be used in conjunction with a stereoscope, where one image is displayed to one eye and the other image to the second eye. Alone, both images look like random dot patterns, but when seen through the stereoscope, a three dimensional figure becomes apparent (Julesz, 1971). In order to test Elizabeth's eidetic memory, she viewed these images through a stereoscope, but not at the same time. Instead the presentation of the second image was delayed, such that she had to store the first image in her memory and later join this memorized image with the second one. She claimed that this task was very easy for her, and managed to recognize the hidden figures without difficulty with delays of up to three days (Stromeyer & Psotka, 1970). No further studies on this subject were published and Elizabeth refused to be tested elsewhere, leading some authors to question the reliability of the study in its entirety (Foer, 2011).

1.1.2. Public searches for superior memorizers

In a few instances scientists have used public media to search for superior memorizers. Between 1970 and 1973 John Merritt published small random dot images analogous to the Stromeyer study mentioned above in various articles in the American popular press, asking that readers try the task for themselves and contact him, if they were successful. He was hoping to find other people with an eidetic memory of this strength. He estimated the combined number of readers of the articles to be in the millions. About 30 adults and children did reach out to him over the years and he met and tested about 15 of them, but none of those tested could reproduce the skill under controlled settings (Merritt, 1979).

A more general search for people with a superior memory was broadcasted by BBC radio in England. Ten people who responded were invited to undergo a battery of memory tasks and were then compared to a group of age-matched control subjects (Wilding & Valentine, 1988). None of the ten subjects could demonstrate superior memory performance across the whole range of tasks. Only three demonstrated superior performance in some of the tasks, but even then they did not exceed the performance of the control group by much; instead they were on approximately the same level as the best participants in the control group.

1.1.3. World Memory Championships

“Memory athletes” are competitors in the “World Memory Championships” (WMC) and other memory competitions. The first WMC was held in 1990 in London. Since then this annual event has grown into an international competition with more than 1000 competitors representing more than 40 countries currently listed in the “Memory Sports World Ranking List” (World Memory Sports Council, 2013). The list is supervised by the World Memory Sports Council, the governing body of memory sports. The WMC consists of ten different disciplines. Material to be memorized includes digits, binary digits, playing cards, random words, (fictional) historic and future dates, names and faces and abstract images. Length of memorization varies between 21.19 seconds (fastest time to memorize a deck of playing cards) to one hour in the so-called marathon events. The average competitor’s performances are far superior to the general population. To finish within the top half of the field at the WMC 2012, a competitor had to memorize more than 500 digits within an hour, and in the correct order. The memory sports records represent the limits of human performance in encoding speed. The current world record for the one hour number memorization event is 2660 digits, correctly memorized in order by the Chinese competitor Wang Feng. A German competitor, Johannes Mallow, managed to memorize 500 digits within just five minutes (World Memory Sports Council, 2013). However, when tested on material to which they are unaccustomed, memory athletes fail to show any superiority (Maguire, Valentine, et al., 2003). This is also true for both autobiographical memory as well as prospective memory (Konrad, 2013).

The first study on memory athletes was conducted immediately after the inaugural World Memory Championships of 1990, by Wilding and Valentine. These researchers used the same test battery that was employed for the group of people, who responded to the radio search a few years earlier, with the addition of some extra tasks designed to allow for comparison of memory for digits with the older, single case studies. All seven participants in the first WMC took part, and three additional people were also tested, who had joined the WMC as spectators and did well during an audience test at the event. Not all of these athletes outperformed the control group, but five of them were at least one standard deviation above the control group on average across all tasks involving immediate recall. The best performances seen in immediate recall were in the tasks in which words, names and telephone numbers had to be memorized, where several of the athletes were more than three standard deviations above the performance of the control group. These three tasks have similar characteristics to several of the tasks used within the competition, and are the most amenable to the employment of specific techniques. In a delayed recall condition with recall occurring one week after memorization, the athletes were not as markedly superior to the controls but nonetheless four of them were more than 1.6 standard deviations above the expected means (Wilding & Valentine, 1994).

In a second study on memory athletes, competitors from the WMC 2000 were studied with regard to their intellectual abilities and structural or functional differences as indexed by brain imaging paradigms, as compared to a control group (Maguire, Valentine, et al., 2003). The authors found that the memory athletes were not exceptional in their general cognitive ability tested as assessed by NART (National Adult Reading Test, verbal IQ, athletes mean 111 ± 8.31) and the WASI Matrix Reasoning Subscale (non-verbal fluid reasoning; athletes mean 12.90 ± 1.79). Instead, their score were in the high-average range. Significant differences were found in memory measures including memory for stories, digit span and a subjective memory questionnaire. All of the memory athletes reported the use of mnemonic strategies; in particular, nine out of ten reported using of the "method of loci".

Using brain imaging, Maguire and colleagues found no structural differences between memory athletes and the control group, but they did find differences in brain activation in regions associated with memory and spatial navigation during memory tasks. A full discussion of these imaging results will be given later in this thesis.

1.1.4. Pi-memorization champions

A different form of memory competition exists for individuals competing for records in the memorization of decimal places of the mathematical constant Pi. As an irrational number, Pi's number of decimal places is infinite, and the places calculated to date go into the trillions. The current World Record for memorizing the most digits of Pi stands at 67890 digits, and is held by the Chinese Chao Lu. In contrast to the athletes competing at the WMC, his digit span is not superior when random digits are given at a pace of one per second, but he only shows superior results when he can set the pace himself (Hu, Ericsson, Yang, & Lu, 2009). Similar results have been reported in one of his predecessors, Tomoyori Hideaki, who held the Pi record from 1987 to 1995. He was superior to controls in self-paced digit memorization, but when memorizing words or stories, he did not excel at all (Takahashi, Shimizu, Saito, & Tomoyori, 2006). Similar to the WMC athletes, both Pi champions used visual mnemonics (Lu e.g. the method of loci), and both reported extensive training and thousands of hours of memorization time for the digits of Pi that they knew.

Another previous Pi champion is Rajan Mahadevan, who was extensively studied by Thompson and colleagues (C. P. Thompson, Cowan, & Frieman, 1993; C. P. Thompson, Cowan, Frieman, & Mahadevan, 1991). An initial difference is apparent in recall time: Rajan recalled 31,811 digits at a pace of about 3.5 digits per second, whereas Tomoyori and Lu recalled less than a digit per second. Similarly to the WMC athletes and in contrast to the other Pi champions, Rajan's digit span was vastly superior in both visual and auditory presentation modes. While he reports having learnt an encoding strategy, it does not seem to include encoding digits into images or referring to known information.

1.1.5. Superior memory achieved by training

The first study to follow the progress of an individual to achieve a superior memory has been published in 1980 (Ericsson et al., 1980) with further details on the case published in follow-up reports (Chase & Ericsson, 1981, 1982). Subject SF, a psychology undergraduate student, and one of three subjects who initially joined the study, is described as a student of average intelligence and memory capacity as compared to other students. He had to train on the digit span task for about one hour a day, several days per week. The digit span task is a classical task to test working memory capacity. Digits are read out at a pace of one digit per second and have to be recalled immediately. The number of items is increased when recall is perfect or decreased when a subject makes too many mistakes. In the version used to train SF, the number of items was adjusted by one after every trial. Average performance on this task is seven digits, and the range for ordinary people is from five to nine digits (Miller, 1956). SF started at exactly this level of seven digits, but he steadily increased his performance. At the end of the study, after more than a year and a half and a total of about 190 hours of training, he had increased his digit span to a level of almost 80 digits (Ericsson et al., 1980). While SF did not receive any instruction on memory improvement techniques, his reports showed that he developed his own strategies, which were comparable to the more formal mnemonic systems. He utilized prior knowledge on numbers which he already possessed, for example track and field scores such as running times (in his spare time SF was an avid runner), or historical and birth dates and clustered references to these in blocks of three or four items. Subsequently, some authors have challenged the opinion that SF was an average student starting with an average memory, since the amount of numbers he had already stored in his memory and used for the task as well as the passion and endurance he put into the study seemed rather unusual (Valentine & Wilding, 1997). However, another of Ericsson´s subjects went on to practice for more than three years and achieved a peak digit span of 104 digits (Richman, Staszewski, & Simon, 1995) and in a near-replication two German students achieved digit spans of 80 and 90 digits, with either a reduced number of possible items or a slower presentation rate, via prolonged training over several months and using mnemonics (Kliegl et al., 1987).

1.1.6. Other forms of memory superiority

There are types of memory performance that appear to be superior but do not in fact meet all of the necessary criteria. One such type is the rapid memory performances demonstrated for example by the Spaniard Ramón Campayo, who achieved memory records in disciplines such as “one second number memorization” (current record: 20 digits), which fail the third criterion of long-term retention. Another is the highly superior autobiographical memory (HSAM) displayed by certain individuals (Parker, Cahill, & McGaugh, 2006), where these subjects show highly unusual long-term-retention of autobiographical events, as well as calendar knowledge and knowledge of important news events that happened during their life-time, but fail to show superiority in controlled memory tasks with limited presentation times (Leport et al., 2012).

1.2. Superior memory via the use of strategies, as compared to natural talent

An important question when studying superior memorizers concerns whether a superior memory can be achieved through endeavor or whether it is an innate talent one simply possesses or does not. Memory athletes unanimously report the use of mnemonic strategies and training as the reason for their performances, but those remain limited to specific tasks; other forms of superiority might instead be inborn.

1.2.1. Skilled Memory Theory

Ericsson (2003) argues strongly for the proposition that almost everyone can achieve a superior memory, and further that all superior memory performances can be explained by the use of good encoding strategies and extended practice. In his “Skilled Memory Theory” Ericsson and colleagues define three principles via which to explain superior memory performances in a broad range of expertise contexts (Chase & Ericsson, 1981, 1982; Ericsson & Staszewski, 1989), as follows:

- Meaningful encoding
- Retrieval structures
- Speed-up by practice

The first principle of meaningful encoding states that superior memorizers encode information via meaningful associations with preexisting knowledge stored in semantic memory. For Ericsson’s student SF, associations were made with running times and dates representing numbers. For a user of mnemonic techniques numbers are encoded by images, where for example S. had visual associations with numbers based on his strong synesthesia.

The second principle postulates that superior memorizers can access memories better because they associate them with retrieval cues in long-term memory at the time of encoding. Whereas in normal memory later access to stored information is a problem, superior memorizers make use of these retrieval cues. A structured set of retrieval cues is called a retrieval structure. SF grouped items by four and named each group. These groups served as a retrieval structure. S. and the memory athletes made use of the method of loci, either with preparation and training in the technique or by informal use of spatial locations along known locations.

The third principle simply states that meaningful encoding and planned use of retrieval structures while memorizing can both be sped up markedly by ongoing deliberate practice. SF and the memory athletes achieved this via scheduled training of their memorization techniques, S. via frequent and sometimes undesired application of his way to memorize.

1.2.2. Naturalists

Wilding and Valentine (1997) agree that Ericsson´s theory is able to explain most of the superior memory performances that have been reported, in particular those where subjects report the use of memory strategies. It is usually the case with these subjects that one finds that the memory skill is limited to known material. Wilding and Valentine do however argue against the generalizability of the theory, and state in opposition to Ericsson that some cases still cannot be explained by the theory alone. Self-reports of people claiming not to use any techniques might be unsatisfactory for various reasons. For example the temptation to present oneself as gifted during stage performances may result in concealment of the fact that the true basis of the performances is mnemonics. There is also the possibility an individual may be using recognized techniques "by accident", without external help or instruction, as seems to be the case with Luria´s subject S (Luria, 1968).

However some individuals for example in their study on the participants who took part in the first WMC, and also in some of the earlier single case studies, did not show the markedly superior performance in singular tasks as did the memory athletes, but on the other hand they had a much broader range of memory skills, applicable to a wide range of tasks (Wilding & Valentine, 1994). Others, e.g. some of the non-strategy users among the WMC competitors in the early 1990`s study, did not even perform above average in immediate testing. In other words, they did not have extraordinary encoding speed, but instead excelled in delayed recall, with almost no forgetting occurring over a week or even longer periods of time. Also Luria reports near-perfect recall by S for a string of digits learned years before in a retest witch S conducted without forewarning. Based on this evidence, Wilding and Valentine suggest that a distinction may be drawn between the strategists, who make use of mnemonic techniques, and the naturalists, who do not. The generalizability of a superior memorizer's skill, as well as their long-term retention capabilities, is used as indicator to decide in which of the two categories a superior memorizer belongs. By way of a specific example, Pi-Champion Rajan is considered to be someone with highly superior performance in both immediate and delayed recall, but neither uses strategies nor chunking the way that SF did; thus Rajan could be an example of an extreme naturalist, thereby contradicting Ericsson´s theory. Ericsson responded to this after testing Rajan himself (Ericsson, Delaney, Weaver, & Mahadevan, 2004). The authors replicated the original study´s (C. P. Thompson et al., 1993) finding on Rajan´s memory skill for both digits and how he grouped numbers, including grouping up to 15 digits and self-reports lacked any references to mnemonic associations. However, when they looked at the recall time of digits within one group of ten digits in a cued recall test, Rajan was faster for the first digits within the group than he was for the later ones. He also made fewer errors at the beginning of a group than for digits in the second half of a group. These findings contradict the assumption that Rajan stores all ten digits of one group in ten different locations and thereby has a superior basic memory capacity. Instead Ericsson et al. (2004) suggest that he makes associations and connections between sets of digits within each block and therefore does not store more than the expected number of chunks. While it remains unusual that Rajan does not make use of artificial or spontaneous associations of digits to other material, the authors suggest that by memorizing digits of Pi for hundreds or thousands of hours he changed his inner representation of digits and can access blocks of digits in long-term memory. Therefore meaningful encoding occurs through the transformation of single digits from a string of digits into these blocks. Since there is no doubt that Rajan uses retrieval structures, he reports forming super-groups of digits and sometimes mnemonically transforms the beginning digits of several subsequent blocks, his performance fits under the Skilled Memory Theory umbrella and suggests an acquired rather than innate skill.

1.3. Mnemonic strategies

Strategies to improve one´s memory have been in use at least since the time of the Ancient Greeks, and have varied in publicity and popularity across epochs. The use of various “artificial” techniques for memorizing has been described as “Art of Memory” or “Ars Memorativa”; this term remained commonplace in historical analyses of such techniques, in particular because of the popular book “Art of Memory” by Francis Yates (Yates, 1966). Those working with memory techniques in education and training more frequently use the name “mnemonic techniques” (also mnemotechnics in American English) with no apparent differentiation between those terms, where “mnemonic” simply stands for any form of memory aid. While playing only a minor role in formal education, the amount of self-help books on memory improvement is vast, and most books base their training strategies on the same fundamental memory techniques (Lieury & Herbst, 2013).

The efficacy of mnemonic strategies has been assessed in different fields including e.g. cognitive psychology, pedagogics, developmental psychology and gerontology.

1.3.1. Visual Imagery

Principles common to the various mnemonic techniques include the use of associations with existing knowledge and an emphasis on the usefulness of visual imagery (Higbee, 1979). In fact, “Rhetorica ad Herennium” (engl. “Rhetoric: For Herennius”, circa 86 BC), the oldest Latin text on rhetoric in existence, mentions the principle of visual imagery as a mnemonic tool. Scientific evidence for a memory bias towards images over words was first provided in the 19th century (Kirkpatrick, 1894). This basic principle alone can be used for example to enhance memory in word-pair tasks, where the mnemonic instruction is to visualize both words together (Bower, 1970a). A common suggestion is that images should be bizarre and unusual, because these tend to benefit memory the most (Yates, 1966). Some studies found a bizarreness effect when testing mental imagery as a mnemonic tool (Merry, 1980; O’Brien & Wolford, 1982), however others did not (Hauck, Walsh, & Kroll, 1976; Wollen, Weber, & Lowry, 1972). One suggestion to explain this discrepancy is that the underlying effect is not the bizarreness of the images, but the distinctiveness (Mcdaniel & Einstein, 1986), or the higher interactivity, of the images (Kroll, Schepeler, & Angin, 1986) created. While visual imagery plays an important role, it is not the only factor contributing to the outcome of mnemonic strategies (Bellezza, 1981).

1.3.2. Keyword mnemonic

The most studied application of mnemonics is the “Keyword” mnemonic. It is mostly used to learn vocabulary but also to study facts and definitions. The basic idea is to first find a word that sounds similar to the word that has to be memorized, or reminds one of it in another way. In the second step this “keyword” should be visualized in an interactive manner together with the translation (Atkinson, 1974).

For example, if one wishes to learn the German word “Erinnerung”, meaning memory, then one could use the keyword “inner room” and visualize an inner room in the brain, where one puts all his memories. To learn that Berlin is the capital of Germany, one could visualize a “bear” for Berlin, either walking through Germany (if one has an image for Germany in mind), or sneezing, where an illness is often predicated on having “germs”.

This method has proven effective in vocabulary learning in a range of settings (Pressley, Levin, & Delaney, 1982), even before the actual term was coined (Ott, Butler, Blake, & Ball, 1973). It has also been shown to work not just for similar languages (some had argued that it is possible to find fitting keywords for languages with a lot of common history such as German and English but not for more distinct language pairs), but also for languages from different cultural background, for example for English native speakers learning Russian (Atkinson, 1974), Chinese native speakers learning English (Bird & Jacobs, 1999) or Malaysian natives learning Arabic (Yaakub & Bakir, 2007). Other studies found that the usefulness might be reduced, when the subjects have to make up keywords themselves rather than using provided keywords (Hall, Wilson, & Patterson, 1981), or in instances where the recall is delayed (Wang & Thomas, 1995). However, neither of these contestations are confirmed across published studies (Lawson & Hogben, 1998; Pressley et al., 1982). Recent studies have found that the way in which subjects are instructed and trained in using the keyword mnemonic (Wyra, Lawson, & Hungi, 2007), and the vividness of the image for the words to be learned (Campos, Amor, & González, 2004), are also important factors. Further peer-generated keywords might be better than experimenter-generated keywords (Campos et al., 2004). Even a gender effect was found in one study (Tabatabaei & Hejazi, 2011).

In summary, the general usefulness of this method of language learning is accepted. It is also applicable to other learning disciplines such as historical, political or geographical details (Pressley et al., 1982), or even university-level neuroscience (Richmond, Carney, & Levin, 2011) and has been shown to be applicable to a range of subjects, including children (Pressley et al., 1982), and even people with some forms of learning disability (Scruggs, Mastropieri, Berkeley, & Marshak, 2010). It can be combined with other beneficial learning strategies, such as retrieval practice, to increase the success of the outcome (Fritz & Morris, 2007).

1.3.3. Face-Name mnemonic

A related mnemonic to the keyword is the face-name mnemonic. It is recommended that it be used when memorizing people's names. Remembering names is very difficult for many (Cohen & Faulkner, 1986), and when people are asked in which area they would like to improve their memory, names and faces is the most common answer given (Higbee, 2001). Here, the instruction to remember names has an effect: When people are given the same word, for example "Baker", and are asked to learn it in association with a given face, their performance is worse when they are told that it is the person's name versus when they are told it is the person´s occupation (Cohen, 1990; McWeeny, Young, Hay, & Ellis, 1987).

Similarly to the keyword mnemonic, an imaginable word is used to remind the learner of the name. For example, for the name "Miller" the mnemonic image might be the activity associated with the profession which the name comes from; for "Bush" therefore one might think of a plant. This image is then visualized together with the person. A common suggestion here is to focus on a characteristic feature of a to-be-remembered face and link the image for the name with that feature. Others suggest visualizing the whole person doing an activity that is related to the image (Konrad, 2013).

The face-name mnemonic has been shown to work in a laboratory setting, where people had to memorize the names of people who they saw in photographs (McCarty, 1980; Morris, Jones, & Hampson, 1978). It also works with caricatures where a characteristic feature is drawn prominently, but not better than with normal photographs (Carney, Levin, & Stackhouse, 1997). It has been shown to be useful regardless of age (Yesavage & Rose, 1984). Besides names, the face-name mnemonic can also be used to associate painters to images (Carney & Levin, 1994), and other facts about people such as their occupation or political opinion, in addition to their name (Carney & Levin, 2012). On the other hand, in some studies the face-name-mnemonic did not work as well in real-life situations such as conversations (Patton, 1994) or a party (Morris, Fritz, Jackson, Nichol, & Roberts, 2005). Reasons for this might include reduced willingness to apply an unused strategy when facing real people, or the high cognitive demand necessary to come up with images and associations that cannot be realized while having a conversation.

1.3.4. Story mnemonic

The story mnemonic is used to memorize a list of items in order. The idea here is to make up a story connecting the items. For abstract information, a mediator is used, similar to how images are made up in the keyword method. An example might be to memorize a shopping list. If the list consisted of milk, potatoes, wheat flour and paper towels, the learner could visualize putting milk in a pot, than adding potatoes; wheat starts growing out of the potatoes which he then wraps in paper towels.

By associating each item with the following, one can follow through a long series of items as long as the first item and the beginning of the story are remembered. Studies have shown the story mnemonic to be effective for various populations including students (Herrmann, 1987) and seniors (Drevenstedt & Bellezza, 1993; Hill, Allen, & McWhorter, 1991). It also works for long lists (Bellezza, Six, & Phillips, 1992; Hu & Ericsson, 2012) and lead to strong between group differences. One study reports the story mnemonic to lead to six to seven times more words remembered in delayed recall of word lists (Bower & Clark, 1969). It has been shown that subjects improve in the story mnemonic through training and that some individuals use this kind of memorization spontaneously without prior introduction (Wenger & Payne, 1995). On the downside, interference can limit usefulness when items repeat within a list (Wenger & Payne, 1995) and study time per item increases with list length (Hu & Ericsson, 2012). The story mnemonic is also termed the “link method” (Massen & Vaterrodt-Plünnecke, 2006).

1.3.5. Peg word mnemonic

The peg word mnemonic is another method used to enhance recall for serial lists. In preparation the learner has to memorize a list of peg words. This is achieved by using words that are easy to remember. A common suggestion is to use a list of words that rhyme with the digits from one to ten, such as “one is bun, two is shoe” (Worthen & Hunt, 2011) and so on. These words are then used as anchors for the items on the lists. Associations are made using visual imagery. Using the example from the previous sub-chapter, for the first item “milk“ one might imagine a bun being soaked in milk and potatoes being put in shoes.

The peg-word method has been shown to be effective for a range of learners encompassing various ages and ability levels, for increasing serial recall (Bugelski, Kidd, & Segmen, 1968; Mastropieri, Scruggs, & Levin, 1985), and also improving performance in delayed testing (Wang & Thomas, 2000).

1.3.6. Phonetic mnemonic and digits

The phonetic system, also known as the “Major System”, is a more complex mnemonic technique aimed at improving one’s ability to memorize digits. It originates from the 17th century (Hrees, 1986; Voigt, 2001). Each digit is associated with a consonant sound and similar-sounding letters. Different memory trainers set up the letters in different ways; a common version used for European languages is based on the table of Aime Paris (1825).

number	sound / letters
0	s, z
1	t,d
2	n
3	m
4	r
5	l
6	ch, j, g, (German sch)
7	k, c, g
8	f,v,w
9	p,b

Table 1: Phonetic code for number memorization based on Aime Paris (1825).

Vowels do not get associated with a number. Using the system, every word has a distinct number it can be encoded into. In the other direction, every string of digits can be translated into a list of words that represent these digits. For example the word memory has the number code 334. For 1492 one finds "turbine" as a possible word. Someone interested in history could visualize the fictional image of Christopher Columbus attaching a turbine to his ship, thus memorizing the year he discovered America.

The amount of preparation required to apply this mnemonic is greater than with other mnemonics. The code alone does not allow for the rapid memorization of digits, since one needs a significant amount of time to come up with images representing the digits. Memory athletes or Pi memorizers therefore prepare tables of 100 or even 1000 images, representing all possible combinations of two or three digits and memorize those (Konrad & Dresler, 2007, 2010) Combined with the method of loci (see below), these tables enable memorizers to perform at a level not attainable for normal learners (Hu et al., 2009; Maguire, Valentine, et al., 2003; C. P. Thompson et al., 1993; Valentine & Wilding, 1997). When tested on unskilled learners, some studies failed to find beneficial effects (Patton, 1986), whereas others did (Bruce & Clemons, 1982; Morris & Greer, 1984). One relevant factor here appears to be whether the participants were supplied with the images or had to made them up themselves (Patton & Lantzy, 1987). Some argue that the phonetic mnemonic only benefits skilled learners with strong cognitive abilities (Hill, Campbell, Foxley, & Lindsay, 1997; Lieury & Herbst, 2013). Contrary to this belief, other studies, which allowed for sufficient training, found that novices can acquire an exceptional memory for digits using the phonetic mnemonic (Higbee, 1997; Kliegl et al., 1987). Despite this, however, even experts on mnemonics seem to be skeptical rather than optimistic about the benefits of the phonetic system (Worthen & Hunt, 2011). In addition to memorizing digits, a prepared table of 100 images corresponding to the phonetic mnemonic can serve as a peg-list for serial list memorization, as with the peg word mnemonic.

1.3.7. Method of loci

The most prominent artificial learning system is the method of loci (MoL). It goes back to ancient times (Hrees, 1986; Yates, 1966), however its popularity has varied over the ages. The basic idea is to prepare a set of locations (Latin: loci) that one can visually walk along in front of the inner eye. Usually, routes along well-known places are suggested for this purpose. Various names exist for such a set of locations including a “route”, a “journey” or a “memory palace”. The method of loci is the most important tool for memory athletes (Ericsson, 2003; Foer, 2011; Konrad & Dresler, 2007; Konrad, 2013; Maguire, Valentine, et al., 2003). It has also been shown to be effective for list learning in the laboratory (Bower, 1970b; Roediger, 1980). Some studies indicate that it might be more effective for younger and healthy subjects, rather than for older or cognitively impaired subjects (Baltes & Kliegl, 1992; Canellopoulou & Richardson, 1998; Kliegl, Smith, & Baltes, 1989; Nyberg et al., 2003;

Verhaeghen & Marcoen, 1996). However, older participants engaged in method of loci trainings do improve, and even show long-term benefits and transfer to everyday memory tasks (Anschutz, Camp, Markley, & Kramer, 1987; Bottiroli, Cavallini, & Vecchi, 2008).

One reason why some studies have failed to show improvements in memory tasks for senior might be due to older subjects having less steep learning curves when training in the techniques, and also the fact that they exhibit some reluctance in actually using the instructed mnemonics (Brehmer et al., 2008; Brehmer, Li, Müller, Oertzen, & Lindenberger, 2007). Despite this, even elderly persons with mild cognitive impairment (MCI) seem to be able to profit from training in an adapted version of the method of loci and related mnemonics (Belleville et al., 2006; Rapp, Brenes, & Marsh, 2002; Troyer, Murphy, Anderson, Moscovitch, & Craik, 2008).

Studies concerned with variations of how the method of loci is presented report it to be more efficient with subject-generated locations rather than experimenter generated locations (Moe & De Beni, 2005), further suggesting outdoor locations should be used in preference to indoor locations (Massen, Vaterrodt-Plünnecke, Krings, & Hilbig, 2009). Studies also show that it works in virtual environments (Legge, Madan, Ng, & Caplan, 2012). Importantly, the lists of locations can be reused without reducing effectiveness (Massen & Vaterrodt-Plünnecke, 2006).

1.4. Memory Training

Most of the studies mentioned in chapter 1.2 deal with individuals with existing memory skills. They might have achieved their skills by training, but nevertheless possessed a high degree of aptitude when first encountered by scientists. There are few studies where individuals achieved superior performance by training under the observation of scientists. On the contrary, most subjects in the studies of mnemonics presented in chapter 1.3 had no prior knowledge of memory strategies and only had a brief introduction and little to no training in the strategies prior to their post-introduction assessment. Higbee (1997) terms the first group “mnemonists”, the second group of individuals training over a prolonged time “apprentices” and the third group with little training “novices”. He called for more apprentice studies, where subjects practice the memory skill beyond their initial introduction, and added his own in which three out of four students managed to memorize a matrix of 50 digits within three minutes and to recall it without any subsequent errors after a total of about 40 hours of practice undertaken over a three-month period. Despite his contestation that apprentice studies allow for the gathering of subjects with more controlled and protocolled training (thereby allowing for more insights than studies in mnemonists), no further studies pertaining to the development of a mnemonic skill have since been published.

Within the last few years memory training has gained a lot of popularity in a slightly different field. As opposed to studying mnemonics, the study of subjects who engage in working memory training by practicing working memory tasks, allegedly without applying strategies, became popular following studies showing the possibilities that exist for the training of working memory (Olesen, Westerberg, & Klingberg, 2004) and in particular following a highly impactful study released in 2008. Jaeggi and her colleagues found significant improvements in fluid intelligence (Gf) subsequent to training in a complex working memory task, namely the n-back task, where Gf improvements correlated with amount of time spent training (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). In just five years the paper received well over 250 scientific citations (273 citations, ISIS Web of Knowledge, September 23, 2013) and sparked a heated debate. Failed replication attempts (Chooi & Thompson, 2012; Redick et al., 2013; T. W. Thompson et al., 2013), and high-impact, high-n studies which found no such effect in similar tasks (Owen et al., 2010), were at odds with other studies which found so-called “near transfer” to other working memory domains (Klingberg, 2010), in various age groups (Brehmer, Westerberg, & Bäckman, 2012), and also with studies reporting “far transfer” to domains including cognitive control and reading comprehension (Chein & Morrison, 2010). Recent comments and reviews (Melby-Lervåg & Hulme, 2013a; Shipstead, Hicks, & Engle, 2012; Shipstead, Redick, & Engle, 2012) criticize the lack of active control groups and the lack of a theoretical foundation for these possible transfer effects of working memory training, as well as a potential publication bias towards positive findings.

Recently, memory strategy training has been brought back to the attention of memory training researchers, since it leads to applicable memory skills, and has reliable and highly replicable training effects in the trained domains, but nevertheless had been somewhat “forgotten” for a time (Mcdaniel & Bugg, 2012). Despite a possible lack of transfer for these training tasks, they do have a beneficial effect for every-day tasks, even for the elderly, but it remains unclear if this transfers to actual benefits for the peoples’ daily lives (Zelinski, 2012). Based on previous findings pertaining to strong effects of strategy use on working memory task performance (McNamara & Scott, 2001; Turley-Ames, 2003) it is also argued that, at least in some cases, the training outcome of working memory training studies might actually be manifest in the acquisition of better strategies by the subjects (Morrison & Chein, 2011) rather than an improvement in working memory capacity per se.

1.5. Neural correlates of mnemonic strategies

The number of existing studies on the neural correlates of mnemonic strategies is quite small. Some insight can be garnered from several studies related to memory strategies, memory improvement, or superior memory performers. In her seminal study on memory athletes, Maguire et al. (2003) looked into structural brain differences using magnetic resonance imaging (MRI), as well as functional differences using functional MRI (fMRI). In a previous study, her group found greater grey matter volume in the posterior hippocampus of London taxi drivers known for their strong memory of streets and routes (Maguire, Frackowiak, & Frith, 1997) which positively correlated with the number of years these individuals had spent working as a taxi driver (Maguire et al., 2000). However this extra volume was not correlated with navigational expertise per se (Maguire, Spiers, et al., 2003). Since the memory athletes used the method of loci, a spatial mnemonic, and exhibited highly superior memory performance for various tasks, it was assumed that similar or even larger structural brain differences in the hippocampus, and potentially also other brain areas, would be seen. However, using voxel-based morphometry (VBM), no such differences were found (Maguire, Valentine, et al., 2003). In spite of this, looking at fMRI data recorded during the encoding of digits, faces and snowflakes, several areas were more active, or exclusively active, across all tasks for the memory athletes as compared to controls. These regions include the right cerebellum, left medial superior parietal gyrus, bilateral retrosplenial cortex and right posterior hippocampus. In addition, only during the digit memorization task (where the athletes made use of the phonetic mnemonic and the method of loci and showed strongest performance compared to controls), the right cingulate cortex, left fusiform cortex and left posterior inferior frontal sulcus were more activated in athletes versus controls (Maguire, Valentine, et al., 2003).

In a study on a Pi memory champion who used similar mnemonics, medial frontal gyrus and the dorsolateral prefrontal cortex were more activated during recall of pre-memorized digits as compared to counting (Raz et al., 2009) paralleled by deactivation of the default mode network (DMN), which generally shows task-independent decreases during tasks requiring external attention (Greicius & Menon, 2004). When the Pi Champion was encoding new digits, activation was found in motor association areas, midline frontal regions, precuneus, lingual and fusiform gyri and, during early encoding, visual association areas. During later encoding activation moved more towards regions of the dorsolateral prefrontal cortex (DLPFC) (Raz et al., 2009). In the structural MRI scan, the only volumetric difference found compared to controls was in the right subgenual region of the cingulate gyrus. A further single case study has been performed on subject DT, who was also part of the Maguire study from 2003, because he competed in the World Memory Championships 2000 and 2001. DT later got same fame as an author and TV documentary personality claiming to be a prodigious savant, whose memory talent is based on unusual ability rather than training. Scientists confirmed synesthesia and Asperger syndrome in DT (Baron-Cohen et al., 2007), but his claims remain disputed with many speculating that his memory performances are based on mnemonic training comparable to other memory athletes (Foer, 2011). One study investigated DT´s brain activation while encoding either structured or unstructured series of digits using fMRI, and found that activity did not differ between such sequences, in line with there being no performance difference. Controls had higher activation in the lateral PFC

(LPFC). Compared to the controls over all sequences, DT had more activation in the bilateral LPFC (Bor, Billington, & Baron-Cohen, 2007).

These studies were concerned with individuals with existing mnemonic skills exhibiting highly superior performances. A different approach is to teach mnemonics to subjects and observe subsequent performance differences due to strategy use. One such study employed young, healthy subjects, who memorized sequences of ten images during an fMRI scan before and after an introduction to, and a very brief training in, the method of loci. Recall performance was significantly improved by the training and fMRI revealed increased activation during encoding in the right inferior frontal gyrus, bilateral middle frontal gyrus, left fusiform gyrus, and bilateral lingual gyrus/posterior cingulate gyrus. During recall after instruction, and as compared to the baseline condition before instruction, left parahippocampal gyrus/retrosplenial cortex/cingulate gyrus/lingual gyrus, left precuneus, left fusiform gyrus, and right lingual gyrus/cingulate gyrus were more activated (Kondo et al., 2005). Another study required that young and old adults memorize word lists, and gave them a very brief introduction into the method of loci. Even the generation of loci was done while in the scanner (Nyberg et al., 2003). Across all subjects, usage of a memory strategy during encoding was associated with increased activity in the left occipito-parietal cortex and left DLPFC. When looking for activity differences in successful versus unsuccessful strategy use, the left occipito-parietal and retrosplenial cortices were found to be more activated (Nyberg et al., 2003).

A later study asked that older adults practice the method of loci for eight weeks, and then looked for anatomical brain changes. Memory training was associated with better memory performance in a source memory task (serial position of words memorized beforehand) and effects on cortical thickness with a trend that memory training reduced atrophy compared to non-training controls. Cortical thickness changes in the right fusiform and lateral orbitofrontal cortex were positively correlated with the improvement in source memory performance (Engvig et al., 2010). An earlier study testing healthy elderly persons before and after five weeks of training in the method of loci in a serial word list learning task demonstrated strong improvements in the memory task from between five to over 23 words correctly recalled (near ceiling performance). Using magnetic resonance spectroscopy (MRS) an elevation of creatine and choline signals in the hippocampus during recall via the method of loci was also found (Valenzuela et al., 2003). Medial temporal lobe (MTL) dysfunction is often cited as the reason for the memory problems associated with mild cognitive impairment (MCI) patients. When looking specifically for differences in the hippocampus, which is part of the MTL, using fMRI, MCI patients exhibit less activity during encoding and recall as compared to healthy controls. Just three sessions of method of loci training were associated with better memory performances and increased hippocampal activity for MCI patients as compared to a matched-exposure control group, indicating that training can lead to partial restoration of hippocampal functioning (Hampstead, Stringer, Stilla, Giddens, & Sathian, 2012).

Studies not concerned with mnemonic training, but which instead look at variations in strategy use, reveal associations between working memory content organization into higher level chunks and increased prefrontal activity (Bor et al., 2003), and also between mnemonic and mathematic coding strategy use with LPFC activity (Bor & Owen, 2007), and between visual working memory expertise and bilateral dorsolateral prefrontal, posterior parietal, and occipitotemporal cortices activity (Moore, Cohen, & Ranganath, 2006). On the other hand bilateral DLPFC activation is more pronounced in the context of use of a poor rote learning strategy as compared to better strategies (Maestú et al., 2003); in addition a range of cortical activation, including DLPC activation, was found in poor performers in a working memory task (Jaeggi, Buschkuehl, & Etienne, 2007), as compared to a relative paucity of activation changes in the good performers.

2. Study 1: Memory athletes

2.1. Introduction

Individuals with superior memory abilities fascinate their peers and have been in the focus of media, public and also scientific attention for a long time (see Chapter 1.1).

The invention of memorizing as a competitive sport gave these individuals a platform to meet and compete. The competitors, that is, the memory athletes, train in a set of specific disciplines with the aim to memorize as many chunks of information in a given period of time as possible (see Chapter 1.1.3, World Memory Championships).

In a seminal study, Maguire et al. (2003) investigated the brains of ten memory athletes (see Chapter 1.5) from the World Memory Championships 2000 and 2001. Since then, memory competitions increased heavily both in terms of numbers of competitors, internationality, and in respect to the record performances in the various disciplines. After the Maguire study it was questioned that there would ever be further memory athlete studies with larger sample sizes (Ericsson, 2003). However, a performance that made the Top 10 of the World Championships back in 2000 wouldn't even make the Top 200 of the current World Ranking List[1]. Since initiation of the study presented here in 2009, more than half of the World´s Top 50 athletes in memory sports could be recruited.

In the present study, various methods including sleep lab assessment, neuroimaging and behavioral testing, were combined in order to study the abilities and characteristics of the participant's memories and further cognitive domains. A control group was selected with its subjects closely matched for gender, age, handedness, intelligence and, for the women, hormonal status (where this has been found to influence memory (Genzel et al., 2012)). While not all participants were available for all parts of the study, a sufficient number of subjects participated in each study part to allow for reliable group-level statistics. It is

[1] www.world-memory-statistics.com, Statistics from the World Memory Sports Council, as opened on October, 1st 2013.

important to note that without exception all athletes who joined the study solely credited their performances to the use of mnemonic techniques (Chapter 1.3).

The first part of the study aimed to investigate possible differences in the sleep characteristics of memory athletes. It is well accepted that sleep plays an important role in memory consolidation, and it has also been suggested that learning influences subsequent sleep, however with quite varied findings (Diekelmann & Born, 2010). An issue for such studies is that subjects are exposed to a huge input of sensory and other information every day; as such the learning session undertaken in an experimental setting might not increase cognitive demand sufficiently to influence subsequent sleep architecture. Simply increasing the amount of information learned is usually not an option to solve this issue since normal subjects have limited memory capacities and simply cannot memorize more. Memory athletes are able to memorize much more data in a given time frame and are also used to very long learning sessions: e.g. Memory World Championships last for three days, including “marathon” events with one hour of memorization time followed by two hours of recall.

For the sleep part of the present study, memory athletes underwent a series of memory tasks without longer breaks during the last five-hour before going to bed in the sleep laboratory. On another day (random cross-over design) they spend a night in the sleep lab without any prior learning having occurred during the day. Since memory athletes are accustomed to such long learning sessions it might be the case that this actually decreases the influence on the sleep. Therefore also the controls had to perform the same memory tasks, trying their best, so that we could see in what way learning which occurs at the extreme capacity limit of subjects would affect sleep. A more detailed account of the possible influence of sleep on memory and how the amount of sleep spend in various sleep stages and other sleep characteristics might differently influence memory consolidation will be discussed later in the introduction of Study 3 in Chapter 4.1.

The memory tasks were selected to address further properties of the abilities of the memory athletes. A False Memory task was done to investigate their susceptibility to false memories.

Also a Directed Forgetting task was done to see whether or not the participant´s ability is of a general nature or specific to instances in which mnemonic strategy are used. Self-paced memorization tasks for playing cards and personal data of people tested the limits of the athletes in tasks that they were familiar with. For these tasks, recall was attempted on the following day, thereby testing for retention beyond short-term memory duration.

In addition to sleep EEG and behavioral tasks, neuroimaging measures were included to compare a short-term memory task to a long-term memory task. This task intended to test Long-Term Working Memory (LTWM) theory (Ericsson & Kintsch, 1995), which extends Skilled Memory Theory (Chapter 1.2.1). LTWM theory suggests that experts (including but not limited to memory athletes) do make use of long-term memory structures directly during encoding, and further that this happens at speed regular subjects can only encode into working memory. The theory postulates that due to long lasting specialization, experts build networks in long-term memory that work for specific contents in their area of expertise and allow new information to be rapidly included into these networks. These areas then supplement working memory abilities.

Also a set of cognitive tasks beyond memory was done to test the result from the Maguire et al. (2003) study that memory athletes do not show superior abilities beyond memory.

2.2. Methods

2.2.1. Subjects

The very special subject group characterizes the study. All memory athletes with German as first language that were ranked within the Top 50 of the official memory sports World Ranking List in 2009 as provided by the World Memory Sports Council[2] were contacted. Any German memory athlete who reached the Top 50 during the duration of the study, until early 2013, was invited to participate at that stage. In total, out of 29 individuals, 25 agreed to participate in parts of the study. Two individuals declined participation since they had not

[2] www.world-memory-statistics.con

competed actively since 2003. One subject could not be included owing to pregnancy and then she fall out of the Top 50 (she last competed in 2007). One subject failed to reply to any invitation.

In addition to the 25 German-speaking memory athletes, for the subparts of the study not relying on German language abilities, three additional athletes from the Top10 of the World were invited, and joined the study.

In total, 28 memory athletes participated at least in parts of the study. All of these individuals had been ranked at least within the Top 50 of the World Ranking List at the moment of their inclusion in the study. Some of them are still competing while others retired some years ago, but all confirmed their ability to perform memory tasks at a superior level. The study sample includes 8 out of 10 athletes from the current (as of September 2013) Top10 of the World Rankings. All subjects without exception declared that they make use of mnemonic techniques and would not be able to display exceptional memory performances without them.

Out of the 28 athletes included 20 are German, five Austrian and three did not have German as first language. 16 are male, with 12 females. 25 are right-handed and 3 are left-handed. All subjects were at least 18 years old when included with an average age of 29.8 ± 10.3 years.

All subjects did the first part of the Culture Fair Test (CFT 20-R; Weiß, 2006), a measure of fluid intelligence. A control group was created matched for gender, age, handedness and intelligence, with the extra criterion for women of hormonal status (contraception, menstrual cycle). Some participants had more than one control subject, in stances where they took part in the different parts of the study on separate days, but where matching was preserved in any case. Since most of the memory athletes showed high performance in the intelligence test (see results), most of the control subjects were recruited within the Munich chapter of

MENSA[3], the international high IQ society, or other associations of gifted students. This had the additional effect that the control group was also drawn from a group of people characterized by outstanding cognitive ability, and therefore no difference in performance motivation between the groups was expected.

2.2.2. Measures

Cognitive Measures

As mentioned previously, the first part of the CFT 20-R (Weiß, 2006) was used as a measure of fluid intelligence. It is a matrix reasoning test with four subtasks and a total duration of 14 minutes test time. Norm values are provided for individuals aged between 20 to 60 years in ranges of five years. The appropriate table was used for each subject with respect of his/her age on the test day. The norm values provided in the test manual are based on extrapolations of prior test-versions and statistics on age-related decline as it pertains to these kinds of tasks since no full empirical assessment of the CFT-20-R exists.

The "Zahlen-Verbindungs-Test" (ZVT) (Oswald & Roth, 1987) was employed as a brief measure of general cognitive ability and processing speed. The ZVT is a trail-making task that measures mental speed and correlates highly with standard psychometric tests of intelligence. Numbers from 1 to 90 are provided on a sheet of paper and have to be graphically connected in ascending order as fast as possible. The test was performed in a single-admission-version. Four trials were performed and mean scores calculated. Norm-values of the ZVT are given for the age groups 16-20 years and every decade between 21 to 60 years; the appropriate table was used for each subject. Norm values were extrapolated beyond the fastest time given whenever subjects were faster. One subject could not do the ZVT due to physical disability.

As a proxy for highest speed achieved in memorizing, the personal best times by our subjects in the discipline "Speed Cards" within an official memory competition was taken. In

[3] http://www.mensa.de/ - addressing the local Munich email list.

this discipline the athlete has to memorize the order of a shuffled deck of playing cards (52 cards, poker deck) as fast as he can. The memorization time is terminated via a self-timing device. After memorization the athlete has a maximum of five minutes for recall, which is done using a second deck of playing cards that has to be arranged in the same order as the shuffled deck. No mistakes are allowed; otherwise the time is void. Memory championships vary slightly in which order disciplines are undertaken, but in every single championship “Speed Cards” is the final discipline. It is always performed with two trials from which only the better one counts, and it is the only discipline in which the athletes race against time in comparison to other tasks in which participants try to provide as much information as possible within a given time frame. As such “Speed Cards” is the best proxy for maximum performance regarding speed of a memory athlete. The current World Record stands at 21.19 seconds. One subject´s Speed Cards score was discarded because at the time of the competition she was just 13 years old and did not compete as an adult.

Memory tasks

By way of confirmation of the memory ability of the memory athletes, two memory tasks comparable to tasks from memory competitions were done. The first one required the memorization of the fictional personal data of people in order to match it with the correct, corresponding face. Approximately 50 index cards, each with a portrait photo on a neutral background, first name, last name, address, city and job were handed out (see Figure 1). All information was selected from lists of the most common names, most common street names, largest cities and most common jobs in Germany, and was randomly assigned to the photos. No piece of information was repeated. Subjects had 20 minutes to memorize as much of the information as possible. Recall was performed on the next morning, with the portrait photos presented in random order as cues. There was no time limit on recall. Each correct first name, correct last name, correct address, correct city or correct job written next to the face it belonged to was awarded one point. In memory championships subjects have to match first and last names with faces, but no other personal data is memorized.

Figure 1: Example index card with portrait photo and personal data. Approximately 50 such cards were handed out, and subjects had 20 minutes to memorize as much data as possible. Recall was performed the next morning with the portrait photos given in random order.

In the second memory tasks subjects had to memorize the order of as many shuffled decks of playing cards as possible within 20 minutes. 6 decks were supplied, each shuffled individually. Recall was performed the next morning using unshuffled decks, which had to be order-matched to the memorized decks. There was no time limit on recall. If a subject recalled an entire deck, it was counted only when recall was accurate. For the last deck that was recalled only partially, we counted all cards that were at the correct serial position. No control subject attempted to memorize a whole deck or more; therefore, for controls, we always counted the cards recalled at the correct serial position. This task resembles the “card marathon” discipline undertaken in memory championships, aside from the unusual length and timing. Normally, memory athletes perform this task over the course of either 10, 30 or 60 minutes of memorization time, with immediate written recall required directly afterwards.

Directed Forgetting

Subjects underwent a directed forgetting task. A total of 100 words were displayed in the center of a computer screen, one after the other, for two seconds each (black font on white background). After a word was displayed an instruction was given, namely either “Erinnern!” (“Remember!”, green font) or “Vergessen!” (“Forget!”, red font). The instruction was also displayed for two seconds. The subject received the instruction to only remember the items followed by the remember command. 50 words were followed by the remember instruction and 50 by the forget instruction pseudo-randomly distributed across all 100 items (see Figure 2).

At recall, subjects were asked to write down all the words that they could remember on an empty sheet. They received the instruction to write down all words, even those followed by the forget command. There was no time limit for the free recall. During scoring every correct word was counted regardless of position. Once a subject indicated he could not remember more, recognition sheets were handed out containing all 100 items shown as well as 50 distractors. Subjects were asked to mark for each word if it was part of the task and, if yes, if it was a remember word or a forget word.

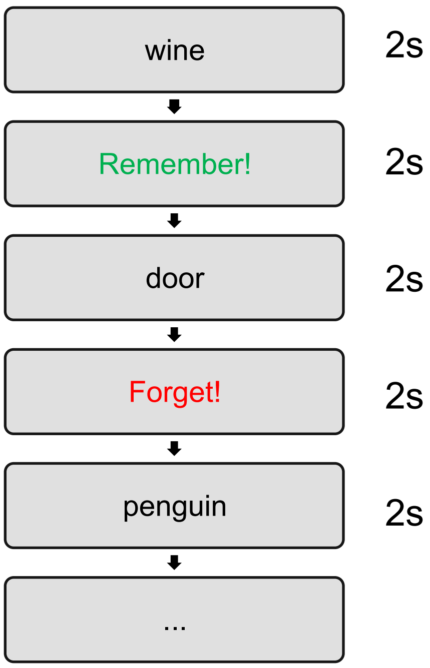


Figure 2: The directed forgetting task. Words (in German, the subject's native language) were displayed one by one for two seconds each followed by an instruction to either remember or forget the item, also displayed for two seconds. A total of 100 words was displayed, 50 for each condition.

The subjects were further split into three subgroups (see Figure 3):

- 9 memory athletes and 9 controls performed recall immediately and another recall on the following morning, after a night in the sleep laboratory,
- 7 memory athletes and 7 controls performed recall only on the following morning, after a night in the sleep laboratory,
- 7 memory athletes and 7 controls, who did not sleep in the sleep laboratory, performed immediate recall only.

For the analysis of the directed forgetting task, first the subgroups with immediate retrieval conditions were pooled (n=16 for both groups) and performance was compared between athletes and controls. Next both next-morning retrieval groups were pooled and the retrieval was compared between both groups. Finally both conditions were compared to look for sleep induced differences. Subjects being tested on both time points were informed about the second recall after initial immediate recall, since assumed future value of items may play a role in sleep consolidation (Wilhelm et al., 2011). Otherwise the group expecting a recall on the next morning would have had benefited from this expectation. This information was provided after the initial recall to prevent rehearsal during recall. Still one has to note that this group saw all items once again during the recognition task.

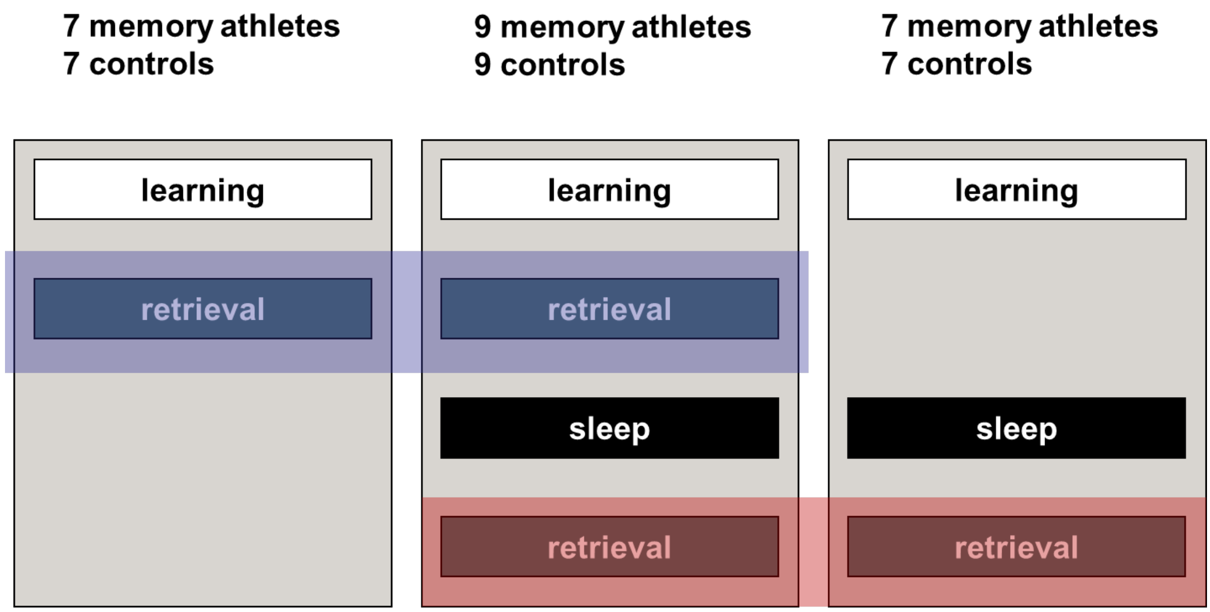


Figure 3: For the analysis of the directed forgetting task, first the groups with immediate retrieval were pooled, next the groups with next morning retrieval were pooled and finally these groups were compared. Note that subjects retrieving at both time points were unaware of a second retrieval during immediate recall but informed about it after the first finish recall to prevent differences in expectancy to those only recalling the next morning.

False Memories

A DRM paradigm was used as a false memories task (Roediger & McDermott, 1995). 18 word lists selected from the original paper were chosen and translated into German. Each list contains 15 semantically related words that are themed around a critical lure, which is associated with all words of a list but is not part of the presented words itself. For example “sugar”, “candy”, “bitter”, “honey” and “girl” are presented, but the critical lure “sweet” is not (Roediger & McDermott, 1995).

Words were presented acoustically, read out by the experimenter at a pace of 1.5s per word. There was a small break of 10 seconds after each list. The recall was done as recognition task with 54 words from the lists, 36 distractors and the 18 critical lures. Subjects had to mark whether or not they had seen an item and how sure there were with their judgment. Regarding recall at different times, the same subgroups had been built as for directed forgetting task.

fMRI tasks

Three tasks have been done during fMRI scanning. Task presentation was projected onto a MRI compatible screen, which the subjects saw via a mirror attached to the head-coil. Paradigms were programmed in Presentation by Neurobehavioral Systems Inc.[4]. 15 memory athletes and 15 control subjects performed the tasks in the fMRI scanner.

The first task during fMRI was a memory task based on binary digits, i.e. sequences of zeroes and ones. In memory competitions memorizing binary digits is one of the disciplines. The memory athletes got a set of 120 binary digits three days before the study day and where asked to memorize those instantly with the aim to still remember them when in the lab. All of the memory athletes who took part in this study had a personal best of more than 300 binaries in the five-minute binary digits memorizing task achieved in official competitions, hence memorizing 120 binary digits was not a difficult task for them. In contrast, memorizing 120 binaries is nothing untrained controls could manage in any reasonable amount of time and therefore the binary task was done with the memory athletes only. Instead, controls had to perform a digit span task with a cumulatively same number of digits to ensure comparable information exposure during the whole learning session. Consultation with athletes revealed that most athletes system is to transform three binaries into a decimal and two decimals into an image and therefore a multiple of six digits was preferred for list length.

In the scanner, the task was split into three parts done in succession.

- Recall of the previously learned 120 binaries in blocks of 24 and motor control task (see below),
- Memorizing of 120 new binaries in blocks of 24 with immediate recall afterwards followed by motor control task,
- Memorizing of 30 new binaries in blocks of 6 with immediate recall afterwards followed by motor control task.

[4] http://www.neurobs.com/

The blocks were designed as follows:

- Memorize: In condition b) subjects first memorized the digits. Five sets of 24 binaries were generated randomly. After the command “lernen” (learn) binary digits were displayed one after the other for 0.5 seconds per digit with an interstimulus interval of 1.2 seconds where the screen remained black. After each series of 24 binaries recall of that series and a motor control task followed before the next binaries were learned. In condition c) the same design was followed except only series of 6 binaries were used and the subjects were explicitly asked to only use working memory and do not purposely apply any mnemonic techniques.
- Recall: Digits were recalled individually in sets of 24. Each block started with a cueing period of the first 12 digits. After command “cue”, the first 12 digits of the blocks were shown one after each other on the screen for 0.5 seconds per digit with 1.2 seconds in between two digits. Subjects were asked to press the button corresponding to each digit shown, using the left button for a “0” and the right button for a “1”. After 12 digits were cued, the command “erinnern” (recall) was displayed and on the following screens an “X“ was presented instead of the digits. Subjects had to recall the correct digits from memory and indicate it with the respective button. For condition c) all six digits were cued in the cue condition and had to be recalled twice in the recall condition to have an equal amount of digits recalled as for the longer series.
- Motor Control Task: After each recall block a control task was done. After the command “drücken” (press button) subjects saw a series of 12 arrows randomly pointing either to the left or to the right (< or >) for 0.5 seconds and had to press the corresponding button. Interstimulus interval once again was 1.2 seconds.

The design of the binary memory task is given in Figure 4.

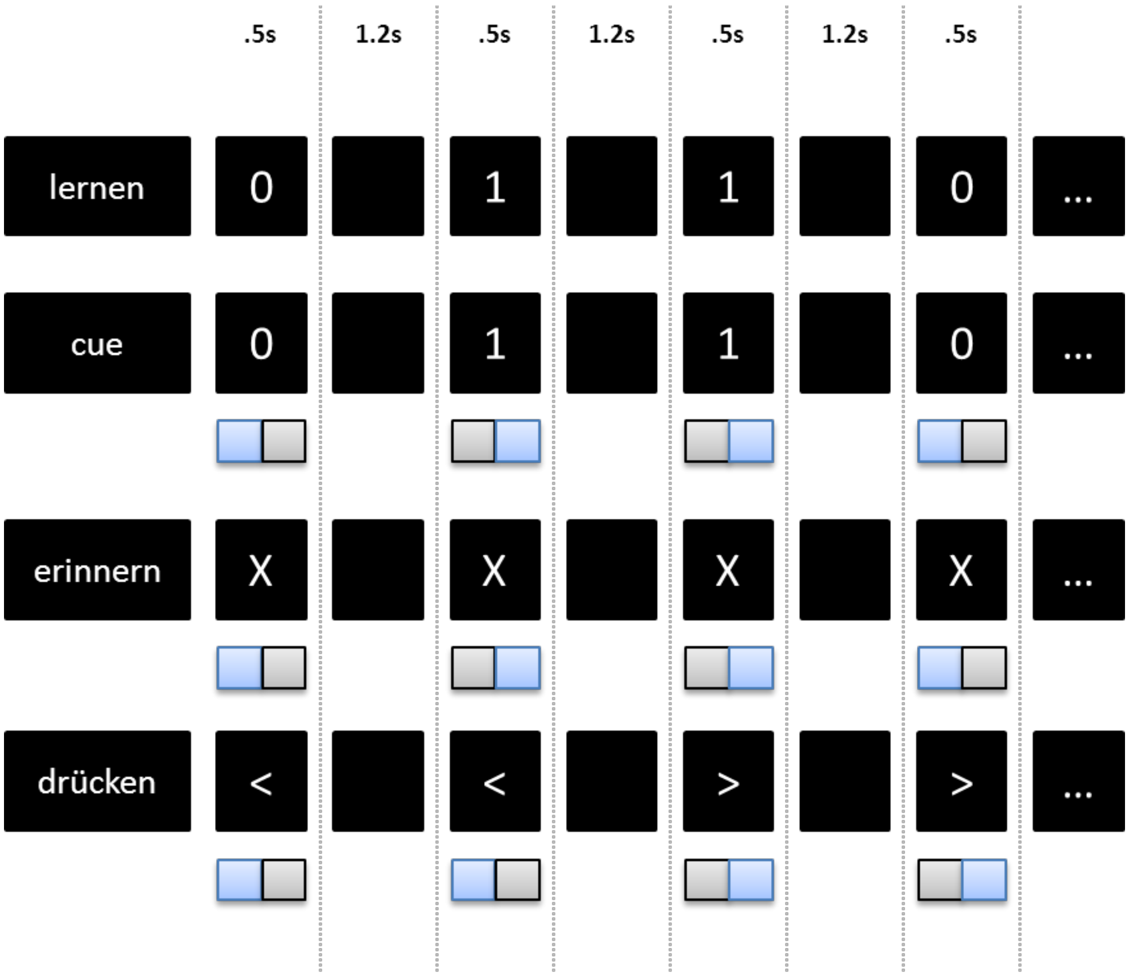


Figure 4: Schema of the binary-memory task in the scanner. During learning (“lernen”) subjects memorized either blocks of 24 binaries (condition B) or blocks of 6 binaries (condition C). The binaries for condition A had been memorized at home three days before the session. During “cue” subjects saw some of the digits they had memorized as a memory cue and pressed the buttons correspondingly. Subjects pressed the left button for a “0”, and the right button for a “1”. During remembering (“erinnern”) only the letter “X” was displayed and subjects pressed the button based on the digits that they remembered. During the motor control task (“drücken”) subjects pressed the button indicated by the arrow.

The second task during fMRI scanning was an n-back task. It is a classical working memory task used within a high number of existing studies and its neuronal correlates are well established (Owen, McMillan, Laird, & Bullmore, 2005). For the present study, a version based on letters as memory content was used. The letters were shown in white font on a black background on the screen one after each other for half a second per letter. After each letter a black screen was shown for two seconds before the next letter appears. N represents a number and was varied between 2, 3 and 4. The subject´s task is to keep the series of letters in working memory and compare them n steps back. If the current letter and the letter presented n before are the same, the subject has to press a button (see Figure 5). The subject´s answer is recorded until one second past disappearance of the letter. Three blocks of each n with 16 letters in each block were run. As a control setting, a task was used where the subject had to press the button when the currently displayed letter equals one shown at the beginning of the block (n=0). In this control condition no memory updating is necessary and only one letter has to be stored in working memory.

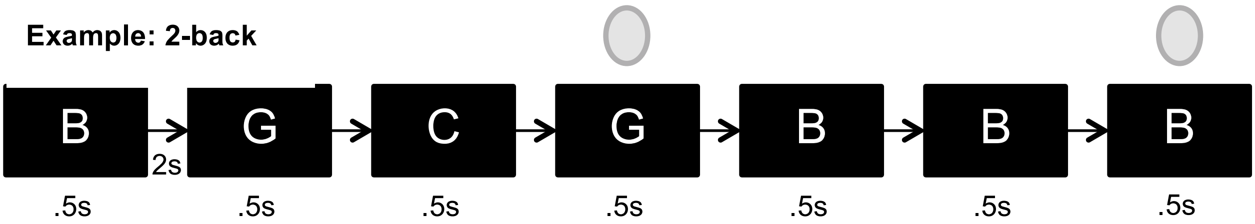


Figure 5: Scheme of the n-back task for an example with n= 2. Letters were shown one after each other for .5 seconds per letter and with an interval of two seconds between two letters. The subject has to indicate with a button press when the current letter is the same is the letter shown n steps before.

The third task in the MRI was an auditory memory task where subjects memorized digits, abstract words with low semantic content (like and, or, in, up etc.) and short stories the heard via MRI compatible headphones. Analysis of the third task is not part of this thesis and will be discussed elsewhere.

2.2.3. Sleep assessment

16 memory athletes and 16 control subjects spend three nights in the sleep laboratory with polysomnographic sleep recording. The first night was not recorded and served as a habituation night. The next two nights were done in random succession as either experimental night, i.e. the night after several hours of memory tasks, or as control night, i.e. night after a day without any memory tasks, where the athletes were also asked to refrain from any training and all subjects to refrain from any learning for studies or else. On the

study day before the experimental night, the subjects tried to memorize more than 1,000 pieces of information over the different tasks mentioned above (compare Figure 6).

Equally in all three nights the subjects slept the whole night in the sleep laboratory with light out between 11pm and 7am. Polysomnography was done using EOG (both eyes, four electrodes), submental chin EMG (three electrodes), ECG and EEG (21 electrodes, 10-20 system, sampled at 250 Hz) and applied by professional sleep lab technicians. Sleep scoring was done by professional sleep scorers based on the Rechtschaffen and Kales criteria (Rechtschaffen & Kales, 1968) using Brainlab Software (Schwarzer Medizintechnik)[5]. For further analysis, sleep scored as sleep stage 3 or sleep stage 4 was combined to SWS.

In addition to sleep stage scoring, full analysis of various sleep characteristics including sleep spindles, rapid eye movements have been done that are not part of this thesis and will be reported elsewhere. This thesis only includes details on time spent in the different sleep stages.

After sleep subjects had time to shower and got served a breakfast before retests started.

[5] http://www.schwarzer.net/medizintechnik/

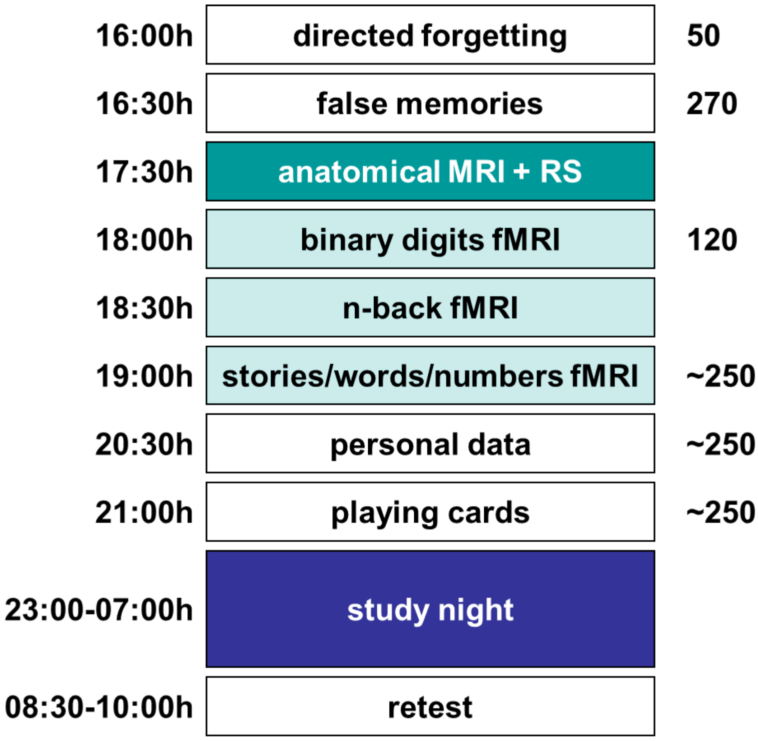


Figure 6: Course of the sleep after learning day for the memory athletes (n=16) and matched controls (n=16) that were part of the sleep-study part of Study 1.

2.2.4. Data analysis

Analysis of behavioral and sleep data was done using SPSS 18. Data is reported as mean ± standard deviation (s.d.) despite where otherwise noted. Significance was assumed for an alpha of 5% and is reported in steps of * = $p < .05$, ** = $p < .005$ and *** = $p < .001$.

Group differences in the memory and cognitive tasks, directed forgetting and false memories tasks were tested using independent samples t-tests with one-tailed significance levels where differences were to be expected (memory tasks) and two-tailed significance tests where not. Correlations between cognitive tests and memory competition performance were done using two-tailed Pearson correlations.

Performance in the n-back task was analyzed using an analysis of variance (ANOVA) with factors group (athletes and controls) and condition (n-back level of 0, 2, 3, and 4). Sleep data was analyzed using a repeated measures ANOVA with factors night (control, learn) and sleep-stage (S1, S2, SWS, REM, WAKE) and between subjects factors group (athletes, control).

For ANOVAs homogeneity of variances was assessed by Levene's Test of Homogeneity and homogeneity of covariances via Box's test of equality of covariance matrices. The assumption of sphericity was tested via Mauchly's Test for Sphericity and Huynh-Feldt correction was used when violated.

2.2.5. fMRI data acquisition and analysis

fMRI was carried out at 1.5 T (Signa Excite, GE, Milwaukee, USA) using an 8-channel head coil and covering 25 AC-PC oriented slices (4 mm thickness, 0.5 mm gap; 64 × 64 matrix, interleaved echo planar images, TR 2000 ms, TE 30 ms) for both tasks reported (n-back and binary digits). fMRI analysis was done with Matlab2008b[6] and SPM8 software[7]. Coordinates are given in MNI space, the AAL toolbox was used for cluster labeling (Tzourio-Mazoyer et al., 2002).

Preprocessing

The functional MR images were preprocessed using SPM8 and MARSBAR[8]. Preprocessing consisted of the following steps. (1) correction of slice time differences due to interleaved images acquisition, (2) realignment to the first volume using rigid body transformation, (3) normalization to the EPI template in Montreal Neurological Institute (MNI) space, (4) reslicing (voxel resolution 2 × 2 × 2 $mm^{3)}$, (5) spatial smoothing using a Gaussian kernel of 8 mm full-width at half maximum (FWHM) and (6) segmentation in native space resulting in tissue probability maps of grey matter, white matter and CSF. The first five images were discarded after preprocessing to remove non-steady-state effects.

N-back

For the first level analysis of the n-back task a high-pass filter (512s) was applied. Nine nuisance regressors (six parameters describing the rigid body transformation and three

[6] By MathWorks, http://www.mathworks.de/products/matlab/
[7] http://www.fil.ion.ucl.ac.uk/spm
[8] http://marsbar.sourceforge.net/

parameters reflecting global signal variations as derived from the cerebrospinal-fluid (CSF) mask and the white matter mask, and from a deep-CSF ROI obtained via MARSBAR were included in the analysis. Positive contrast images were produced for each n. In the second level analysis these images were included in a full factorial model with factors group (athletes, controls) and condition (n = 0,2,3,4). For effects of condition a statistical maps were created using a voxel-wise family wise error (FWE) correction with a threshold of $p_{FWE} < 0.05$ and a threshold of k ≥ 25 voxels applying t-contrasts ([-1.5,-.5,.5,1.5] for activation associated with, and [1.5,-5,-.5,-1.5] for deactivation associated with increased task difficulty) on the condition (n = 0,2,3,4) to investigate for brain activity changes in both directions associated with increasing task difficulty. For condition x group interaction and main effect of group a threshold of uncorrected $p < 0.001$ per voxel followed by cluster based multiple test correction procedure (FWE, $p < .05$) was used.

Binary

Two athletes did a preliminary version of this task and had to be excluded from the analysis. For the first level analysis of the binary digit memory task a high-pass filter (512s) was applied and the same nine nuisance regressors included as for n-back. Since the tasks were done in individual scanning session in direct succession, separate first level analysis had to be done for previously learned binary digits (condition A), long lists of binaries learned in the scanner (condition B) and short lists of binaries learned in the scanner (condition C). Since the interest was in comparing the three conditions, the contrast recall > motor control was produced as first level. These contrasts were than included in a second level analysis, which was done as full factorial model set up with the factor condition (A,B,C) as dependent measures with unequal variances assumed. Pairwise comparisons were done between the three conditions with a threshold of uncorrected $p < 0.001$ per voxel followed by cluster based multiple test correction procedure (FWE), with significance defined as cluster p-values

< 0.05. A conjunction analysis was done at threshold of uncorrected p < 0.005 with FWE correction at cluster level was done to confirm overlap of clusters.

2.3. Results

2.3.1. Cognitive Measures

		N	mean	s.d.	range
Age (years)	Athletes	28	29.8	10.3	19-58
	Controls	28	29.2	9.6	20-58
ZVT (IQ)	Athletes	27	136.1	15.5	106-165
	Controls	28	123.3	16.4	100-160
CFT (IQ)	Athletes	28	131.5	12.0	103-165
	Controls	28	133.1	12.0	112-158

Table 2: Cognitive test data for the memory athletes and matched controls. One athlete could not perform the ZVT due to physical disability. Data are given as mean ± standard deviation (s.d.), and range.

Performance in the CFT and ZVT is given in Table 2.

		N	mean	s.d.	range
Age (years)	Athletes	28	29.8	10.3	19-58
	Controls	28	29.2	9.6	20-58
ZVT (IQ)	Athletes	27	136.1	15.5	106-165
	Controls	28	123.3	16.4	100-160
CFT (IQ)	Athletes	28	131.5	12.0	103-165
	Controls	28	133.1	12.0	112-158

Both groups were matched by age and IQ and therefore did not differ in those dimensions. However, in the ZVT a significant group difference was found ($t(53) = 2.989$; $p < .005$) with the athletes being faster than the controls. In the fluid intelligence task (CFT), the population norm is given as IQ = 100 with an s.d. of 15. Both groups mean performance was above the second standard deviation of the population norm.

In the discipline Speed Cards, based on the best time achieved in an official memory competition as on the statistics web page of the World Memory Sports Council[9], the average time of the memory athletes was 57.6 ± 22.8 seconds with a range of 21.19 seconds to 90 seconds for the memorization of a shuffled deck of 52 poker cards with perfect recall afterwards.

When correlating Speed Cards performance with ZVT and intelligence, strong correlations were found (see Figure 7). Time in seconds for Speed Cards and ZVT raw scores in seconds correlated significantly ($r = .547$; $p < .005$), Speed Cards times and ZVT scores, age-corrected by the norm value tables giving the various IQ scores correlated significantly ($r = -.642$; $p < .001$). Speed Cards time and fluid intelligence as measured by CFT also correlated significantly ($r = -.432$; $p < .05$). Faster times in Speed Cards and ZVT are indicative of better performance, and therefore times in seconds and IQ scaled-performances show negative correlations.

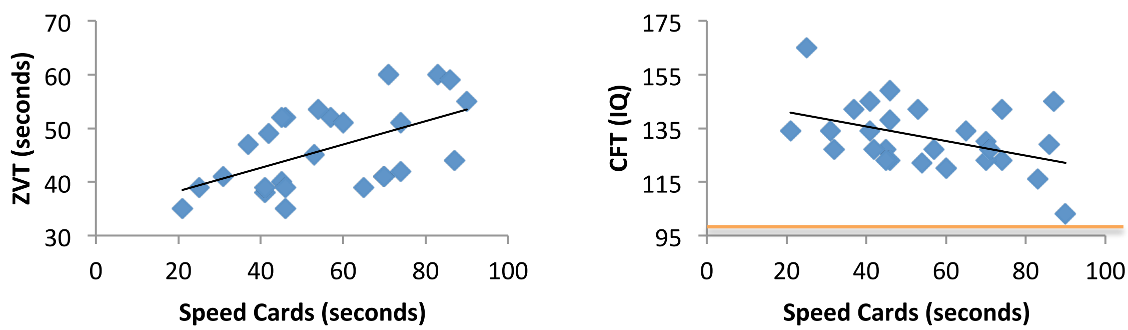


Figure 7: Scatter plots of the athletes performances in the Speed Cards event (as assessed during an official competition) compared to processing speed as measured by ZVT (left panel) and fluid intelligence (Gf) as measured by CFT (right panel) with linear regression lines. For ZVT and Speed Cards, faster times equate to better performances, while Gf negatively correlated with Speed Cards results. The horizontal orange line in the right panel denotes the population norm of IQ = 100. In both cases correlations were significant ($p < .005$ for ZVT; $p < .05$ for Gf).

2.3.2. Memory tasks

[9] www.world-memory-statistics, October 1st 2013.

Playing cards and personal data memory tasks were performed after a series of other cognitive demanding tasks, right before going to bed while getting the EEG attached. Therefore fatigue and distraction might have negatively influenced performance. Cards data is not available for two athletes (resulting n=14), performance of personal data is not available for two controls (resulting n=14) and one athlete (resulting n=15).

As expected, large group differences were found in the pure memory tasks. In the cards memorization task memory athletes on average remembered 170.6 ± 63.7 cards whereas controls remembered 15.6 ± 7.7 cards. Range in athletes was 63 to 300 cards (312 was the maximal possible score), in the controls the range was 6 to 34 cards. The difference was highly significant ($t(28)=9.680$; $p < .001$; Cohen´s $d = 3.4$). In the personal data task, memory athletes on average correctly recalled 84.2 ± 40.6 pieces of information (range 25 to 135) while the controls on average correctly recalled 40.8 ± 15.0 pieces of information (range 18 to 75). The difference was highly significant ($t(27) = 3.763$; $p < .001$; Cohen´s $d = 1.4$).

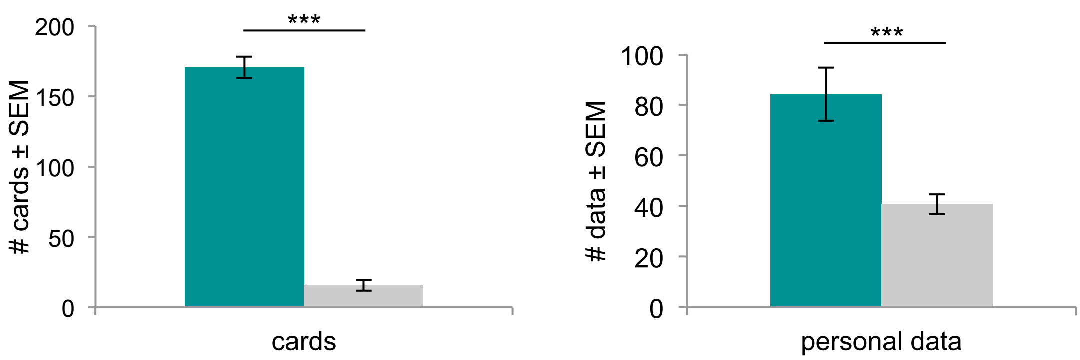


Figure 8: Memory task performance for cards and personal data. As expected, memory athletes highly outperformed the controls in these tasks that were similar to tasks done at memory championships. *** $p < .001$.

2.3.3. Directed Forgetting

First the immediate retrieval of all subjects with immediate retrieval after the task were pooled (n = 16 for athletes as well as controls) and compared between athletes and controls.

In the free recall of words that had been followed by the remember instruction memory athletes could recall far more words than controls: 44.8 ± 5.4 words (athletes) versus 18.1 ±

6.2 words (controls). An independent samples t-test proves this difference to be highly significant ($t(30) = 12.982$; $p < .001$) and the effect size is very large with Cohen´s $d = 4.6$ Seven out of 16 memory athletes but no control correctly recalled more than 95% of the 50 remember items. Since memorizing lists of words is part of the special ability of memory athletes, these findings were expected. Far more interesting is the comparison of words recalled that had been followed by the "forget" instruction. Here the memory athletes still recalled 5.4 ± 3.8 words correctly and the controls 3.6 ± 2.7 words. The t-test revealed this difference not to be significant ($t(30) = 1.494$; $p = .146$).

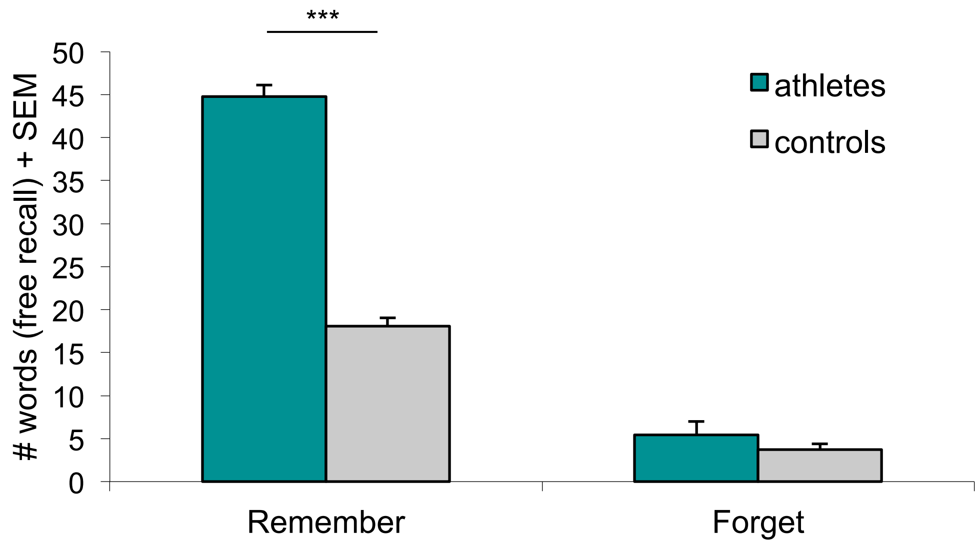


Figure 9: Directed Forgetting, immediate recall (free recall). Number of words correctly recalled in free recall immediately after the learning. n=16 for both groups. Difference for forget items was not significant. *** $p < .001$.

After the free recall, a recognition task was done and performances are given in Figure 10. If just counting the judgment whether or not a word was presented, memory athletes on average recognized 48.9 ± 1.7 remember words out of 50, with many correctly recognizing all 50 remember words. Controls recognized 42.0 ± 6.0 remember words. As in the free recall, this difference was highly significant ($t(30) = 4.459$; $p < 0.001$, Cohen´s $d = 1.6$). Out of the 50 forget words, athletes recognized 33.4 ± 6.7 forget words and controls 28.8 ± 7.2 forget words. This difference showed a non-significant trend ($t(30) = 1.855$; $p = .073$). These performances are given based on recognizing a word, regardless if it was correctly identified as remember item or forget item. If scored specifically for words correctly recognized as forget items, the differences between athletes (30.8 ± 8.2 forget words recognized as such) and controls (20.1 ± 9.7 forget words) turned significant ($t(30) = 3.382$; $p < .005$; Cohen´s d = 1.2), but one has to keep in mind that most memory athletes could correctly identify most or all remember items and thereby any further word recognized could only be from the forget condition.

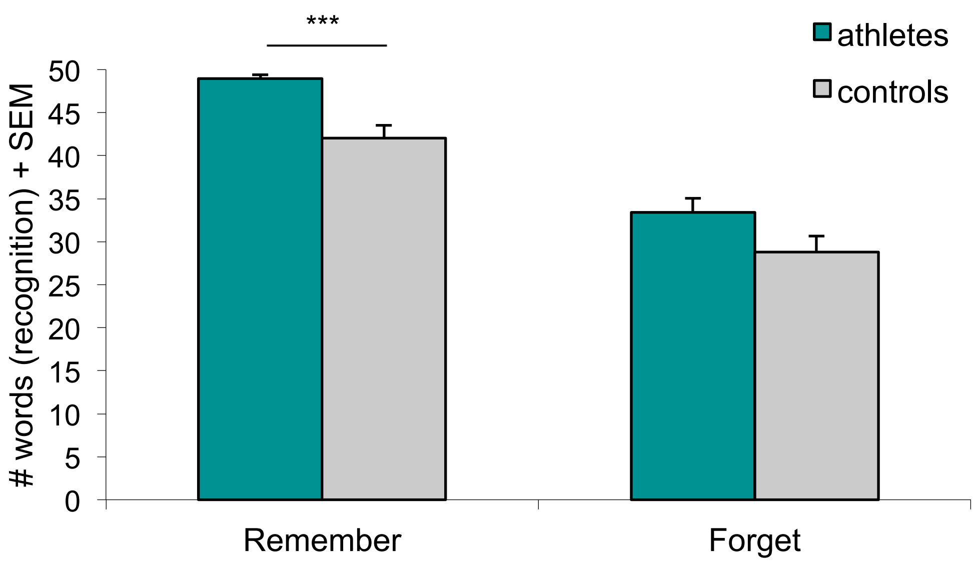


Figure 10: Directed Forgetting, immediate recall (recognition). Number of words recognized out of 50 remember items and 50 forget items. The difference for forget words was not significant. *** $p < .001$.

When looking at the pooled results of both groups recalling on the next morning, the same pattern is found. The group differences in free recall on remember items is highly significant (43.6 ± 8.4 remember words for athletes versus 15.9 ± 10.1 remembers words for controls; $t(30) = 8.398$; $p < .001$; Cohen´s d = 3.0) and as well in the recognition test for remember items (47.9 ± 2.2 vs. 39.6 ± 6.8; $t(29) = 4.503$; $p < .001$) with no recognition data available for one of the athletes. When looking into forget words, the group difference is not significant, neither in free recall (4.2 ± 4.4 vs. 3.0 ± 2.2; $t(30) = .972$; $p = .339$) nor recognition (26.5 ± 9.4 vs. 23.3 ± 7.4; $t(29) = 1.059$; $p = .298$). When looking for items correctly recognized as forget items, the difference turned significant (24.1 ± 9.4 vs. 13.9 ± 9.9; $t(29) = 2.922$; $p < .01$).

When comparing the nine subjects per group that recalled twice (immediately after memorizing and on the next morning) to the seven subjects recalling only on the next morning, there was one single difference striking out: In the athletes group the value for free recall of forgetting items in those recalling twice was 6.6 ± 4.5 forgetting words on the next morning compared to only 1.1 ± 1.2 forgetting words in those recalling on the next morning, a significant difference ($t(14) = 3.091$; $p < .01$). No other difference was significant or just trending in athletes as well as subjects when those recalling twice were compared to those only recalling on the next morning. When comparing the immediate recall performance and the next morning performance in the nine athletes with paired t-tests, no difference in free recall or recognition was found to be significant. Comparing the immediate and delayed recall in the nine controls using paired t-tests, significant forgetting was found via recognition recall in both remember words ($t(8) = -3.203$; $p < .05$) and forget words ($t(8) = -6.457$; $p < .001$), see Figure 11.

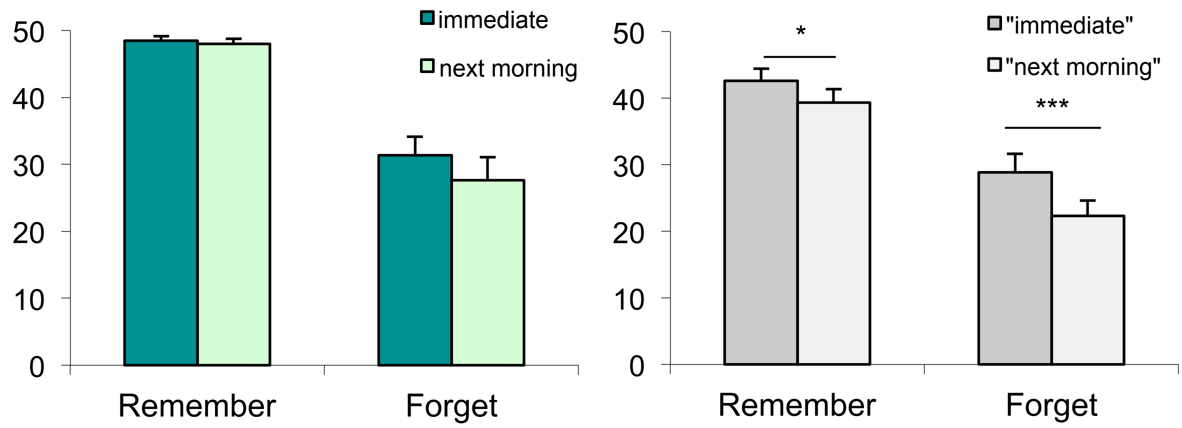


Figure 11: Athletes (left) and controls (right) recall performance (recognition) on immediate recall and next morning recall. Significant forgetting was only found in controls. *** $p < .001$; * $p < .05$.

2.3.4. False memories

Similar to the “remember” condition in the Directed Forgetting task, it was expected that memory athletes remember more items that actually were presented than the controls (n=16 for both groups). First the subjects with immediate recall were pooled and the immediate recall data was compared. The recall was done as recognition task with 54 words from the lists, 36 distractors and the 18 critical lures equaling a total of 108 asked. Out of the 54 actually presented items, memory athletes on average recognized 45.2 ± 5.6 words, whereas controls remembered 38.5 ± 6.7 words. This difference was significant ($t(30) = 3.081$; $p < .01$; Cohen´s $d = 1.1$). For distractors few mistakes were expected in both groups and out of 36 distractors presented athletes wrongly identified 5.5 ± 3.7 words and controls 7.4 ± 4.1 words. This difference was not significant ($t(30) = 1.349$; $p = .187$). Most interesting though were the critical lures, the items that actually were not presented, but triggered by semantically related words that were. Out of 18 critical lures athletes thought to recognize 8.1 ± 4.8 lures and controls 12.4 ± 3.1 lures. This difference was significant ($t(30) = 3.025$; $p < .01$; Cohen´s $d = 1.1$) showing that athletes were less prone to false memories than controls.

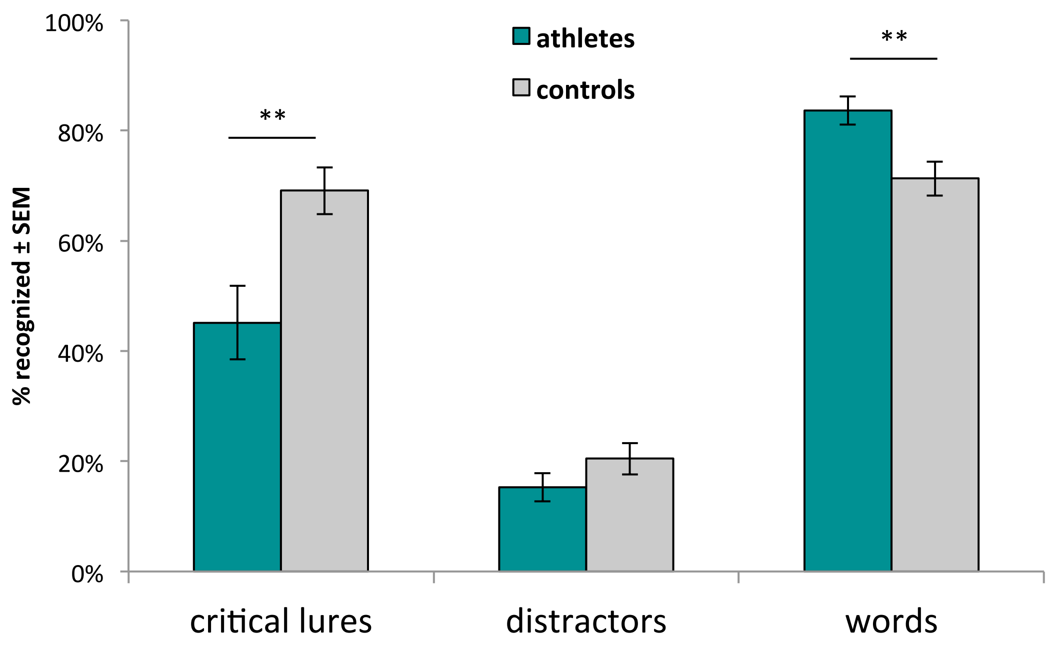


Figure 12: False Memories, Immediate Recall. Recognition performance in the false memory task for athletes (n=16) and controls (n=16) in % of items presented per group. 18 critical lures were set with 15 semantically related words per lure presented. ** $p < .01$.

When looking at all subjects who recalled on the next morning, the same pattern is found. One memory athlete did not follow instructions in the morning recall, therefore $n_{athletes} = 15$ and $n_{controls} = 16$ for these comparisons. As in the immediate recall, athletes recognized significantly more words that had been presented (40.5 ± 8.2 words vs. 34.8 ± 9.0 words;

$t(29) = 1.863$; $p_{one\text{-}tailed} < .05$), athletes were significantly less prone to falsely identify critical lures (7.4 ± 4.3 lures vs.10.6 ± 3.3 lures; $t(29)=2.303$; $p < .05$) but groups did not significantly differ for wrong recognition of distractors (7.9 ± 6.0 distractors vs. 9.4 ± 5.7 distractors).

When only looking into the 9 subjects who did both immediate and next morning recall, both groups showed forgetting for words that actually had been presents. Athletes on average forget 3.2 ± 3.3 words overnight, a significant forgetting at $t(8) = 2.922$; $p < .05$) and controls forget 4.3 ± 4.7 words overnight, which was also found to be significant ($t(8) = 2.772$; $p < .05$). Controls showed a decrease in false memories and falsely identified less critical lures on the next morning than immediately after recall (-2.6 ± 2.4 lures overnight change, significant at $t(8) = 3.261$; $p < .05$). In the athletes the overnight change in critical lures was not significant (+1.1 ± 2.7 lures overnight change, $t(8) = 1.229$; $p = .254$) and the nominal increase in lures was driven by one athlete who went from just one lure identified in immediate recall to ten lures identified on the next morning. The group difference regarding falsely recognized critical lures remained significant nevertheless.

2.3.5. N-back

The n-back task was performed in the MR scanner. Due to technical malfunction the responses of one athlete were not recorded. Therefore for analysis of the n-back task performance $n_{athletes} = 14$ and $n_{controls} = 15$. There were 16 targets each in the 0-back, 2-back and 3-back condition and 13 targets in the 4-back condition distributed over the course of the task. Performance on targets hit is given in Figure 13.

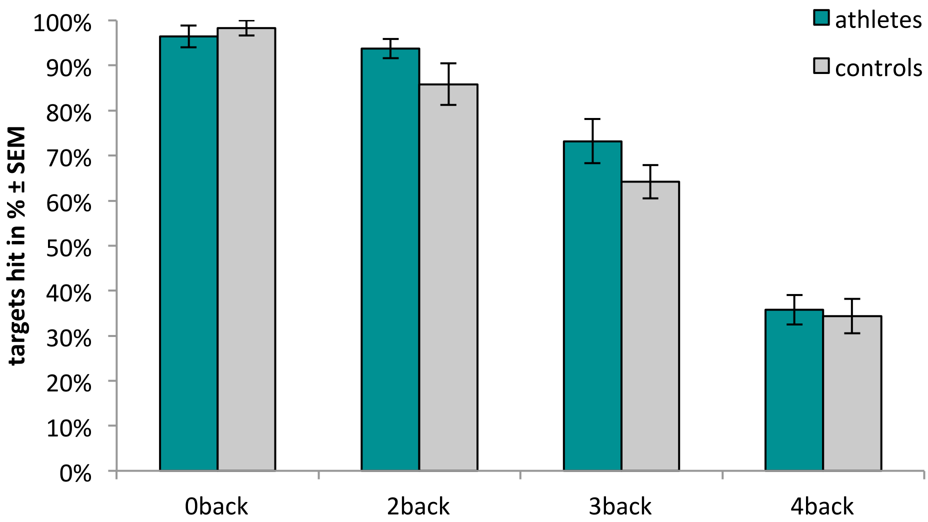


Figure 13: Performance in the n-back task as percentage of targets hit for athletes and controls. In total there were 16 targets for 0-back, 2-back and 3-back and 13 targets for 4-back. There was so significant group difference.

As expected, the main effect of condition was significant ($F(3,81) = 234.143$; $p < .001$; $\eta_p^2 = .897$) with significant differences between each condition and increasing difficulty from 0-back to 4-back.

ANOVA further revealed no significant group x condition interaction ($F(3,81) = 1.719$; $p = .170$; $\eta_p^2 = .060$). The main effect of group was not significant either ($F(1,27) = 1.577$; $p = .220$; $\eta_p^2 = .055$) and there were no differences on targets hit in any of the condition with $t_{0back}(27) = .655$ ($p = .518$), $t_{2back}(27) = 1.514$ ($p = .142$), $t_{3back}(27) = .151$ ($p = .151$) and $t_{4back}(27) = .267$ ($p = .791$).

Looking into reaction times rather than correct hits produced comparable results. No significant group x condition interaction (Huynh-Feldt corrected $F(1.934,52.52.216) = 1.305$; $p = .279$; $\eta_p^2 = .046$), no main effect of group ($F(1,27)=2.154$; $p = 154$; $\eta_p^2 = .074$), highly significant main effect of condition (Huynh-Feldt corrected $F(1.934,52.52.216) = 35.031$; $p < .001$; $\eta_p^2 = .565$) and no differences between groups in any condition were found.

Yet again looking into mistakes made rather than hits showed the same results: No significant group x condition interaction ($F(3,81) = .334$; $p = .801$; $\eta_p^2 = .012$), no main effect of group ($F(1,27)=.064$; p = 803; $\eta_p^2 = .002$), highly significant main effect of condition interaction ($F(3,81) = 29.095$; $p < .001$; $\eta_p^2 = .519$), and no differences between groups in any condition.

2.3.6. Sleep data

A repeated measures ANOVA with factors night (control, learn) and sleep-stage (S1, S2, SWS, REM, WAKE) and between subjects factors group (athletes, control) was done. The night*sleeps-stage*group interaction was not significant (Huynh-Feldt corrected $F(1.907,57.205) = .695$; $p = .497$; $\eta_p^2 = .023$), and neither were the night*group interaction ($F(1,30) = .007$; $p = .933$; $\eta_p^2 = .000$) nor the sleep-stage*group interaction (Huynh-Feldt corrected $F(2.680,80.414) = 1.421$; $p = .231$; $\eta_p^2 = .045$).

When looking for main effects, there were no significant main effects of night ($F(1,30)=.177$; $p = .677$; $\eta_p^2 = .006$) and no significant main effect of group ($F(1,30)=1.017$; $p = .321$; η_p^2 = .033). Only the main effect of sleep-stage was significant (Huynh-Feldt corrected $F(2.680,80.414) = 192.507$; p < 001; $\eta_p^2 = .865$) since of course different amount of time is spent in the various sleep stages.

In summary, the analysis failed to find any differences in time spent in the various sleep stages between both groups or between the two nights. Sleep stage duration did not change for any group due to learning nor does it differ between the groups. Figure 14 and Figure 15 display the comparison of time spent in the sleep stages for both groups and both nights.

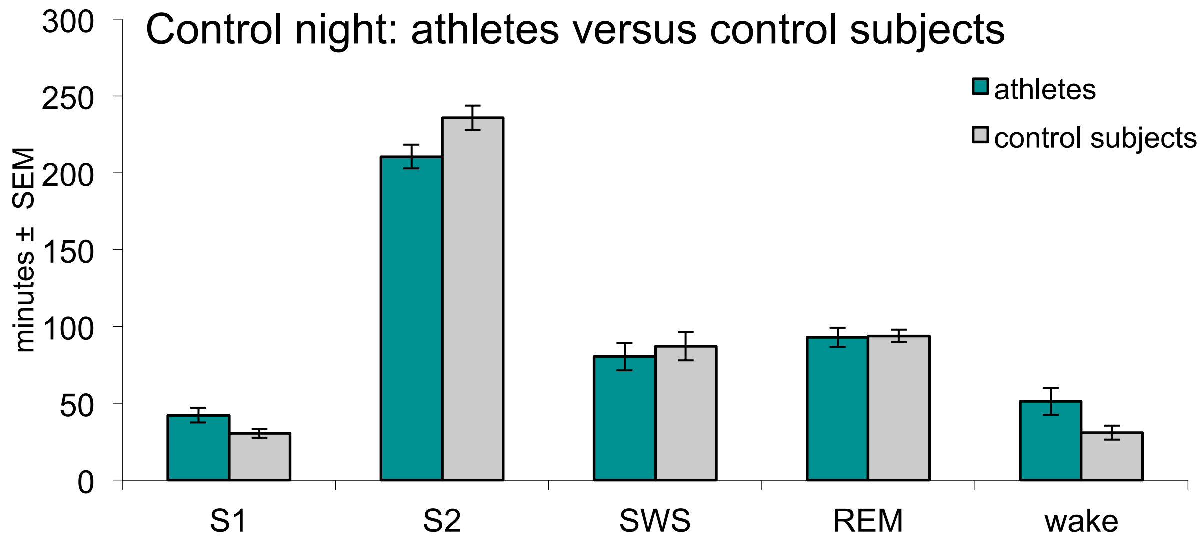
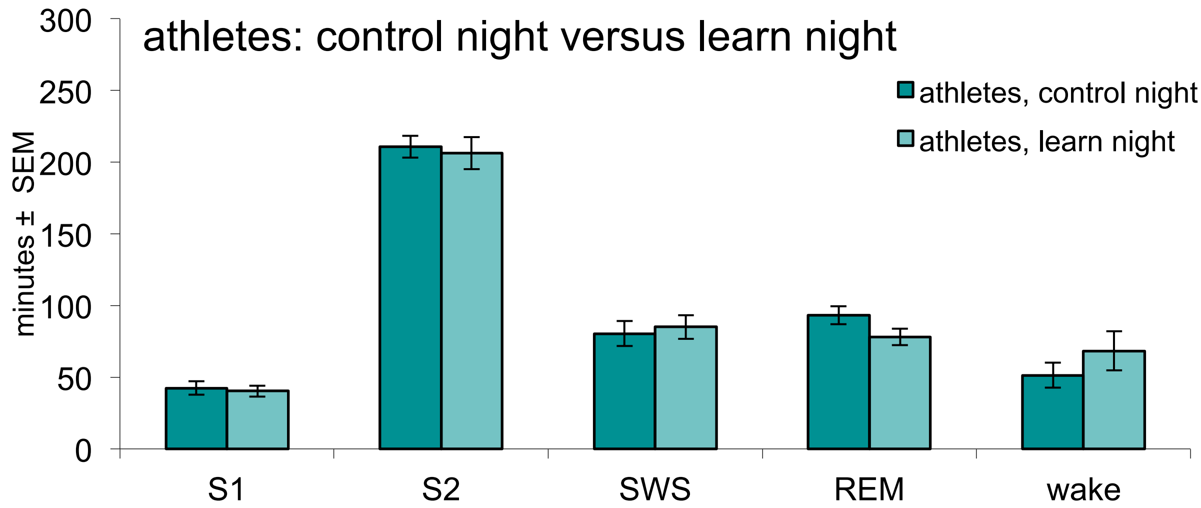


Figure 14: Minutes spent in the different sleep stages by athletes and control subjects in the control night. No significant group differences were found. The same was true for the learning night (not shown),



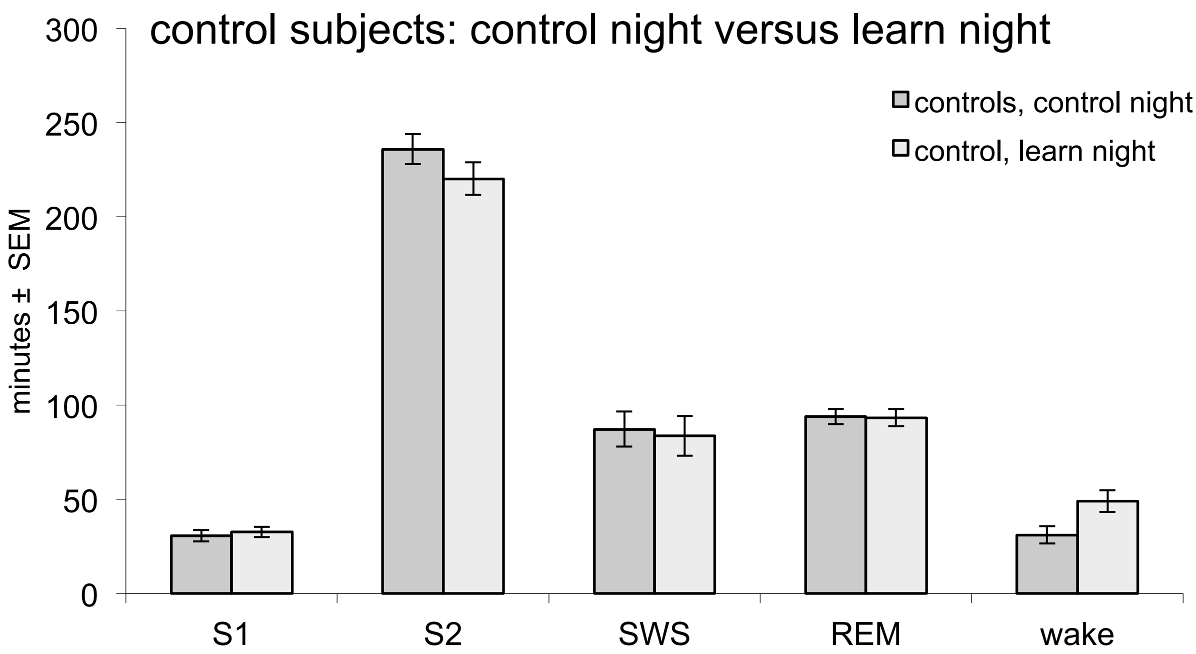


Figure 15: Comparison of the night after learning and the control night for athletes (top) and control subjects (bottom). No influence of intense learning on time spent in the sleep stages was found in either group.

Full analyses of various sleep characteristics have been done and will be published elsewhere (in prep.). In summary, there were also no group differences and no learning influence on number of sleep spindles or other common sleep spindle characteristics, number of rapid eye movements, REM density or any other studies sleep parameters.

2.3.7. fMRI

N-back

The effects of condition for the n-back task (gathered at $p_{FWE} < .05$) revealed a range of activations associated with task performance (see Figure 16) including (MNI coordinates (x,y,z) of peak voxel given): bilateral parietal cortex (-40, 40, 44; 20, -68, 54), bilateral superior frontal cortex / left precentral gyrus (28,0,56), bilateral precuneus / cingulum (-2,-46,30), bilateral medial frontal cortex (-10,38,56), right precentral gyrus (48,8,32) and left temporal cortex (-44,-66,-6) matching well to findings of meta-analysis on the n-back task (Owen et al., 2005). Besides activations, deactivations were found in the negative effect of condition for the Default Mode Network (DMN) as was expected (Hampson, Driesen, Skudlarski, Gore, & Constable, 2006).

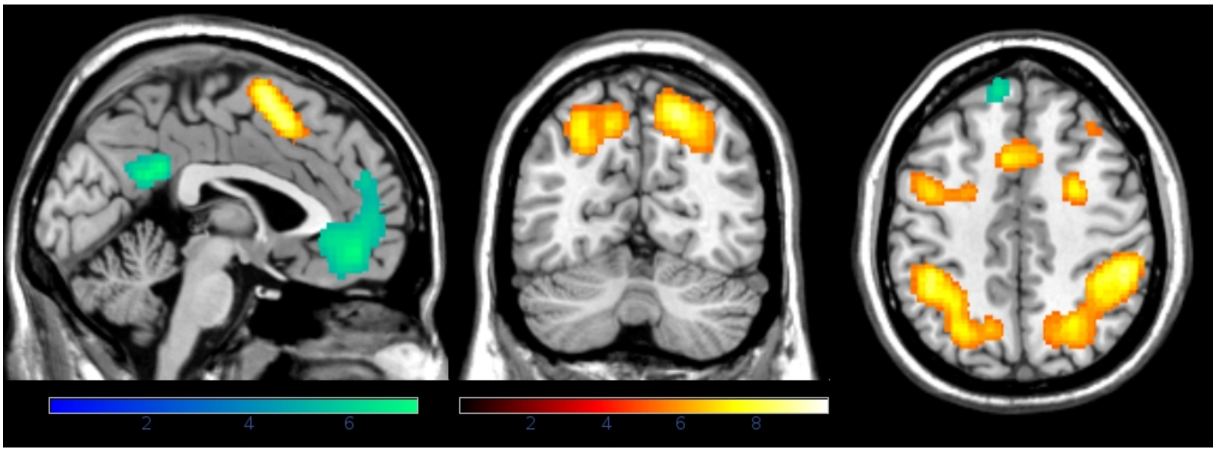


Figure 16: Statistical map for the main effect of condition in the n-back task gathered at voxel-wise corrected $p_{FWE} < .05$. Warm colors represent more activation associated with task difficulty, blue colors deactivation associated with task difficulty.

The main effect of condition with an uncorrected p-value of $p < .05$ was then used as mask when looking for task-related group differences. Sampling at an uncorrected threshold of $p < 0.001$ no significant clusters were found after applying cluster based multiple test correction in both the condition x group interaction as well as the group effect indicating no differences in memory athletes and control subjects in regards of brain areas activated when doing the n-back working memory task.

Binary

The three recall conditions (A: 120 previously learned digits, B: 120 newly learned digits, C: sets of only 6 binaries; in all cases contrasted against the motor control task in the individual's fixed-effects analysis were compared pairwise at a threshold of uncorrected $p < 0.001$, corrected at cluster-level (FWE) with $p < .05$. No significant cluster were found in the contrasts A > B, B > A, A > C and B > C.

Significant differences were found in the C > A contrast in the superior frontal gyrus (-18, 18, 48) and also in the C > B contrast the superior frontal gyrus was found to be more active (-6, 28, 54). These findings are shown in Figure 17.

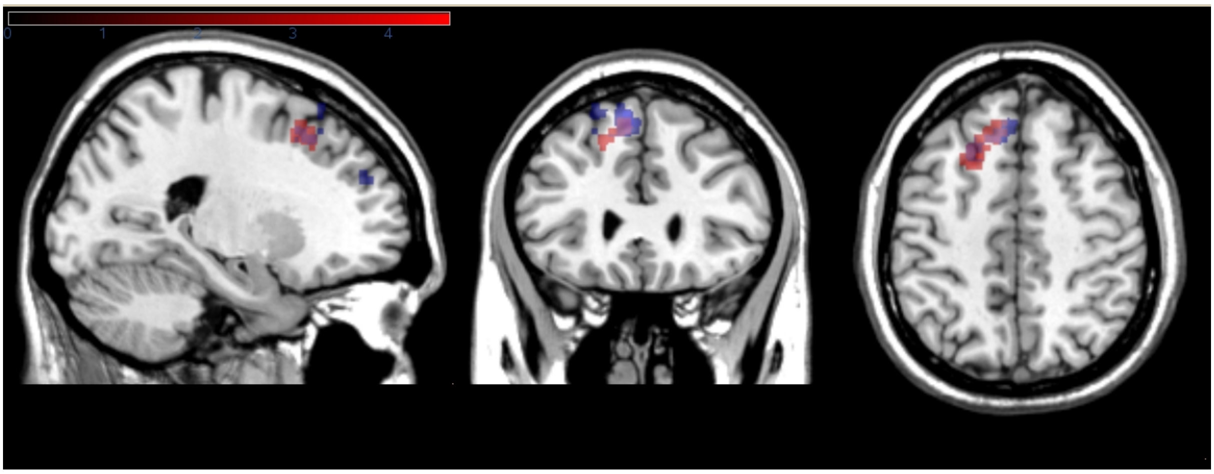


Figure 17: Contrast between the recall from a pure working memory (C) and the recall from long term memory condition (A) shown in blue found a activation in superior frontal gyrus (MNI coordinates (x,y,z): -18, 18, 48) as did the contrast between working memory (C) and recall from the condition in which 120 binaries had just been learned using mnemonics (B) shown in red (-6, 28, 54).

To verify if the overlap of the clusters found by visual inspection of the resulting contrast maps is indeed significant a conjunction analysis was done on the contrasts C>A and C>B with a threshold of $p_{uncorrected} < .005$ at voxel level. In deed the cluster in the superior frontal gyrus (-8,28,54; cluster size of 1263 voxels; see Figure 18) turned out to be significant in this analysis.

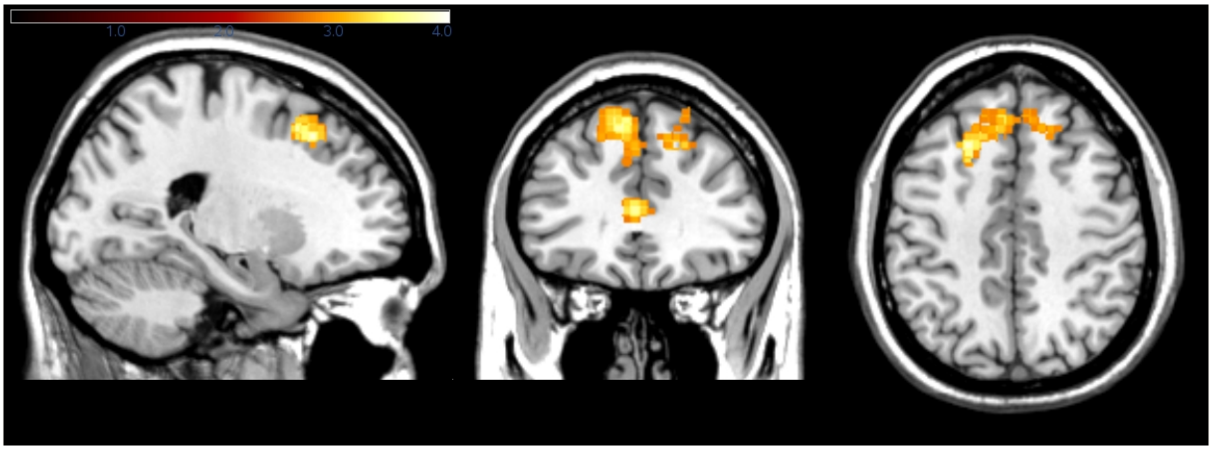


Figure 18: Conjunction analysis of areas active in both contrasts between working memory recall (C) over sets involving 120 binaries (recent (B) and old (A)) sampled at uncorrected $p < .005$ and FWE correction on the cluster level with $p < .05$.

2.4. Discussion

After the study by Maguire et al. (2003), this is the first one to look into memory athletes as a study population. Since 2003 a lot has changed in memory sports. Records and numbers of participants increased a lot. A total of 28 memory athletes could be recruited for participation.

2.4.1. Cognitive Abilities

Maguire and colleagues reported that memory athletes were in the higher average range in tasks of general cognitive ability, namely the matrix reasoning subscale of the Wechsler Abbreviated Scale and the NART test for verbal intelligence. The study presented in this thesis cannot confirm these findings. In the present study participants were tested using the CFT matrix reasoning task and the ZVT, a trail-making task testing processing speed. Highly superior performances in these tests were found, the average IQ being estimate at above 130, two standard deviations above population average by the CFT. Furthermore even correlations between memory sports performance as measured by the best time achieved in the fastest memory sports event Speed Cards and both CFT measured IQ and ZVT were found. One explanation of this discrepancy with the ten-year-old study is found in the increase of performance level at memory championships. Therefore this sample was even more selective than the previous one since level of memory performance that had to be achieved for inclusion in the present study was substantially higher than ten years ago besides the higher number of subjects now included. This also readdresses the question, if such high cognitive abilities are really necessary to achieve memory records. As found in the Maguire et al. (2003) study also all of the athletes in the present study report the use and intense training of mnemonic techniques and credit their performances on them. The same is true for other cases of mnemonists presented in the literature in the last decade (Ericsson et al., 2004; Hu et al., 2009; Raz et al., 2009) with one exception being subject DT (Baron-Cohen et al., 2007), who claims to have a superior memory due to innate ability caused by Asperger syndrome and synesthesia, but is also known to have promoted mnemonic techniques in the past (Foer, 2011). Combined these recent studies provide strong evidence that nowadays memory records can only be achieved with the use of mnemonic strategies and after intense training. This does not rule out the possibility that besides having the best strategies, having high cognitive ability is necessary as well to achieve this performance level. The high IQ (131.5 ± 12.0) average found in the reported sample (note that not a single athlete had a below average IQ) is indicating towards this notation. Factors that also might contribute to this IQ distribution are a bias in people getting in contact with mnemonics and memory sports. Due to the promises of improved learning, it is much more likely to hear about these techniques as a student or when working in academia than when leaving education for more practical professions with less theoretical learning required. People with high cognitive abilities are also more prone to enjoy cognitive demanding activities in their free time and more able to concentrate for prolonged times as necessary when practicing memory sports. Additionally increased IQ and processing speed might be outcomes of the training rather than prerequisites for success. Studies on working memory training suggest that fluid intelligence can be improved by training of just a few weeks (Jaeggi et al., 2008) whereas memory athletes usually trained for years before reaching the top. Since the athletes in the present study varied highly in the number of years they were doing memory sports (2 to over 15 years) and improved their performances over time by continuous training, the strong correlation between processing speed measured by ZVT and memory encoding speed for playing cards is at least an indicator that mnemonic training improves processing speed. If processing speed was a fixed trait, skilled aspiring memory athletes would score high in the ZVT despite still improving in Speed Cards. In general ZVT and IQ are highly correlated constructs (Oswald & Roth, 1987) and strong correlations were also found in the study presented in this thesis. Yet the IQ matched controls showed ZVT performances fitting to the achieved CFT, but the athletes were showing ZVT performances even superior to their already high CFT performances. Since the memory athletes were studied after their training it is not possible to further judge on what is innate and what is achieved by training. To further address this issue, study 2 looks into training of naïve subjects.

2.4.2. Generalizability of memory ability

Maguire et al. (2003) found memory athletes not to excel in visual memory tasks. Wilding and Valentine (1997) similarly reported superior memorizers not to be better in memorizing pictures, faces (but names) or snowflakes. Already Chase and Ericsson (1982) found that huge improvement in a memory skill for digits do not improve memory for letters. Studies on Pi memory champions did not find unusual results in paced memory tasks (Hu et al., 2009; Takahashi et al., 2006). In the present study it could be shown that memory athletes excel on memorizing various kinds of information like digits, binary digits, personal information including names and faces, playing cards and words. They showed superior memory with different forms of presentation (visually, acoustically, paced, self-paced), in different encoding speeds (starting at .5s per item) and durations (a few seconds up to 20 minutes). Superior memory was found with immediate recall after encoding but also when retrieving in the next morning. On the other hand in a working memory task (n-back), where they could not easily apply their method, athletes did not outperform intelligence-matched controls. Very clear is the finding from the directed forgetting task. When asked to memorize 50 words displayed for only two seconds per word, memory athletes can do this really well. Many got all 50 correct and the difference to the control group is more than striking. Yet, in the forget condition when the instruction was not to remember the words, they are on the same level as controls. We strongly asked them to try as much as they could to remember the forget items to rule out that our subjects would not reproduce them due to disliking the instruction. When asked during debriefing, most reported, they only applied the method of loci when the remember instruction came by visualizing the word as an image an associating it to a location. This clearly indicates applying the mnemonics is essential for the performance and despite year-long intensive training they did not acquire a general memory skill but rather a set of methods they can apply on a huge range of tasks.

2.4.3. Mnemonics influence memory processes beyond capacity

Interesting for evaluating theories on memory are some findings, which indicate a different form of memory encoding when mnemonics were applied. In the directed forgetting task athletes were extremely good in the free recall of remember items, but also excelled in the recognition task. Also in the False Memories task, athletes recognized items that actually were presented significantly better than control. Both findings show athletes do not just achieve a better access to memory that is generated anyway, but actually do encode more information. The False Memory tasks´ second finding that memory athletes are less prone to fall for critical lures is also interesting, since it indicates that memory athletes do not just store more information, but also have a higher awareness of memory and somewhat deeper level of processing, always assuming they used the mnemonics on the information. During debriefing they also filled in a questionnaire on strategies used. These were rather informal and elaborateness on the comments varied. Therefore no correlation was possible between performance outcome and strategy use, yet a very interesting details was found when looking into strategies applied in the False Memories task. In this task, words were read out rather rapidly at less than 2s / word. Several athletes noted, they could not stand the pace and had to stop using the method of loci and switch to alternative methods, because they could not make up the images and associations fast enough. Some other athletes reported that they could stick to the method until the end of the task despite the high pace. When splitting the athletes in these two groups, those who switched to alternative methods show nearly equal amounts of false memories as the controls: These participants “recognized” as many critical lures as items presented, which is in line with a study that suggested more intelligent people to be even more prone to fall for false memories. On the contrary, the memory athletes using the method till the end of the task had wrongly identified few critical lures at all.

2.4.4. Long-term Working Memory

Also findings from the binary digit task in the scanner show different memory processing in athletes. When the subjects learned 120 binary digits in the scanner and immediately recalled them, despite the high pace during presentation they did not forget them right away as to be expected for a working memory task (Baddeley, 2003): When asked somewhat later if they still were able to recall the digits, they could easily reproduce them (no formal testing was done on delayed recall, so this remark has to stay anecdotal at this time). Ericsson´s Long-term Working Memory Theory (Ericsson & Kintsch, 1995) provides a theoretical framework to explain for these findings.

It suggests that experts in various fields, by deliberate practice and endurance, achieve the ability to directly store information related to their fields of expertise into long-term memory at a pace normally only short term memory can achieve. Since in contrast to working memory (Miller, 1956) this “long-term working memory” (LTWM) does not show a capacity limit of seven items, much better memory performances can be realized and utilized as extension of regular working memory by expert performers. In that way memory athletes are a special group of experts, since memorizing is their area of expertise. The capacities to directly store information in long-term working memory to did not develop by automation, but were deliberately learned by the subjects. Through deliberate training, memory athletes can address this information as fast as they recall information stored in working memory, as for example seen in the False Memory task (words read out at one per 1.5 seconds) and binary digits task (0.5s per digit). The fMRI findings support the LTWM model. Memory athletes learned binary digits in three different conditions. First they learned a sequence of 120 binaries some days before coming to the lab using their techniques and recalled them in the scanner. Next they memorized 120 additional binaries while being in the scanner and immediately afterwards had to recall those. Finally they were asked to memorize sets of 6 binaries only and refrain from using any mnemonic technique but should just keep those in working memory and recalled those.

The first condition has clearly to be considered a long-term memory task, as the binary digits had been learned days before, well beyond working memory duration (Baddeley, 2003). It is also obvious that the third condition was a pure working memory task as it involved the memorization and immediate recall of just a few items within working memory capacity; it was similar to working memory tasks done in many studies (A. R. A. Conway, Kane, & Al, 2005); and memory athletes were asked and subsequently confirmed not using mnemonic strategies. Contrasting C against A found significant differences in superior frontal regions that have frequently been shown to be important for working memory (Cabeza & Nyberg, 2000; Olesen et al., 2004). In line with the literature, activation differences were found in the left hemisphere only, since a verbal, non-spatial task was used whereas a spatial task would be expected to rather activate the right hemisphere (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000).

The second condition (B) is the most interesting: The digits are learned and immediately recalled and thus have to be considered a short term memory task. One study looked into various memory loads for working memory and also found increase in frontal activity related to the number of items to be held in working memory (Rypma, Prabhakaran, Desmond, Glover, & Gabrieli, 1999). This would suggest even more frontal activity in this condition of the presented study. However, the number of binary digits to be learned clearly exceeds regular working memory capacity (Miller, 1956). One possibility is that converting the digits into images functions as chunking, leading to a more efficient use of the existing working capacity by holding images in working memory that each encodes several binaries.

One study had tested the neural correlates of using of a chunking strategy in a working memory task, which led to a strong increase in working memory performance (Bor et al., 2003) and also found an increase in frontal activity for the chunking items at encoding and no difference between chunking items and regular items at retrieval. A prior study already had shown that integrating different items in working memory makes the task easier and is also associated with increased frontal activity (Prabhakaran et al., 2000). Hence if memory athletes memorize the 120 binaries by better chunking or integrating with locations while holding the data in working memory, this would also suggest an increase in frontal activity. The long-term working memory theory on the contrary would imply more long-term memory regions acting as working memory replacement. This would mean a frontal decrease compared to the working memory task and that the long-term memory condition and the recall of just learned 120 binaries is more similar.

What was found in the present study matches the LTWM prediction. In the second condition frontal areas active during working memory recall were not activated but deactivated as much as in the long-term memory condition whereas no differences at all could be found between the long-term and short-term recall of binaries learned using mnemonics. While it is somewhat surprising that no additional activation was found for the long-term recall over working memory, these findings support the LTWM model.

In addition to the binary task, fMRI findings from the n-back task give evidence that working memory processes in the memory athletes do not differ much as their performance did not differ from controls and no additional brain activation was found in the athletes in the n-back task, while the main effect of condition showed expectable task activation in both groups matching the broad literature on this task (Owen et al., 2005). The debriefing questionnaire showed that no athlete could apply his mnemonic technique during the n-back task, but a few guessed that could adapt their method to the task provided enough time given.

2.4.5. Sleep

Some of the athletes slept in the sleep lab. It was of interest if the high input of information would influence sleep architecture. Sleep plays a role in memory (Maquet, 2001) and it is still discussed if sleep architecture changes due to learning before sleep (Rasch & Born, 2013). An alternative theory is that memory consolidation processes happen in sleep anyhow and recency is just one criteria to mark information as important for reactivation during sleep (Bendor & Wilson, 2012). The presented study rather supports the second notion. Besides more than a thousand pieces of information memorized by the athletes, their sleep architecture did not change at all compared to a control night. One might argue that athletes sleep pattern has adjusted to extreme memory input over time, but that can also be ruled out since sleep between athletes and controls did not differ either and also the memory controls did all the tasks trying to remember as much as possible, and their sleep architecture was not affected either. These findings show that sleep is robust against prior learning and processes during sleep influencing memory do not need to be triggered by learning. Study 3 looks further into this topic.

3. Study 2: Intense mnemonic training

3.1. Introduction

Mnemonic techniques allow for large improvements in memory capacity. However, the amount of training necessary to achieve outstanding results has been considered a limiting factor in experiments concerned with the efficiency of mnemonics. While it remains debatable as to whether 'natural' superior memorizers exist (Valentine & Wilding, 1997) most subjects with superior memory acquired their skill by deliberate training and the use of memory techniques (Ericsson, 2003).

Only few studies have looked into the process of acquiring such memory skills. In a seminal case study, a volunteer started with a normal digit span of 7, which he eventually increased to 80 digits after 20 months of intense training (Ericsson et al., 1980). During the course of his training, he developed semi-systematic methods of relating digit strings to meaningful information related to his own experiences. In follow-ups, and also in a replication study, a few additional individuals achieved similar levels of memory capacity for digits (Chase & Ericsson, 1982; Kliegl et al., 1987; Richman et al., 1995). In one study plus a follow up phase, groups of younger and older adults practiced the method of loci for serial word list recall, but the performance was optimized against speed for few items rather than maximizing number of items stores (Baltes & Kliegl, 1992; Kliegl et al., 1989; Kliegl, Smith, & Baltes, 1990). No other group studies have been published thus far on intense mnemonic training known to the author.

With the rise of memory sports and the World Memory Championships, it became easier to study superior memorizers (Maguire et al., 2003, Chapter 2 of this thesis), but the subject pool is nonetheless still somewhat limited. An alternative research approach to superior memory would therefore be to instruct naive subjects in the use of mnemonic techniques. This approach has the additional advantage that specific mnemonic techniques can be trained, thereby minimizing variance introduced by different mnemonic approaches.

However, most studies teaching mnemonics to naive subjects are limited by the lack of routine in using these techniques (cp. Chapter 1.4). While trained subjects perform better than controls without knowledge of mnemonics, their results are still far below those seen in published cases of superior memorizers. Higbee (1997) calls such subjects "novices" in contrast to people who acquired a superior memory by persistent training under the direction of researchers, who he terms "apprentices". The problem with apprentice-studies is that an intensive, and often month-long, training period is necessary. Thus, frequent drop-outs have to be considered. Such studies are rather expensive and may only produce more single superior memorizers instead of providing larger groups of trained people.

The aim of Study 2 was to assess the chance of acquiring exceptional results after a rather limited training time of about six weeks, following a weekend course in mnemonics in a group of normal subjects compared to a matched waitlist control group. In addition to measures of memory performance, further goals included investigating transfer effects on non-trained cognitive tasks, as well as exploring possible neuronal changes associated with the mnemonic training using neuroimaging techniques. These findings were compared against the findings of Study 1 for discussion, whether memory athletes are different to a normal sample, or alternatively if most people could achieve their performances with equal amounts of training.

3.2. Methods

3.2.1. Subjects and Design

35 healthy, male subjects were included in the study. Inclusion criteria were right-handedness, no history of psychological or neurological diseases, no substance abuse, no current medication, no experience in memory training, German mother tongue, not meeting any exclusion criteria for magnetic resonance scanning and provision of written informed consent for the scientific measurement of their performances and agreement to be included in either group. The MWT-B vocabulary test (Lehrl, Triebig, & Fischer, 1995) was used as a

screening measure for linguistic ability, with a minimum raw score of 18 as cut-off mark. The short depression scale BDI-V (Schmitt et al., 2003), with a maximum score of 35, was deployed as a screening measure for depressive symptoms, and an fMRI safety screening questionnaire was used as a screening measure for physical health and fMRI suitability.

The age range for inclusion was 18 to 30 years and subjects were recruited via mailings and flyers at the various universities in Munich as well as vocational schools and through word-of-mouth. Subjects were randomly allocated to either training or waitlist control groups following the initial session which sought to balance group sizes. Subjects received an honorarium of 200 Euros in the training group or 100 Euros in the control group subsequent to completion of the entire study. Subjects in the control group were offered the chance join a memory course of equal content after completion of the study, as part of their compensation. The study was approved by the local ethics committee at the Ludwig Maximilians University.

Training Group

Figure 19 shows the study design for both groups. All subjects took part in an initial session comprising behavioural testing and the first neuroimaging session, including structural brain scans for screening purposes. After the first session, 20 subjects were assigned to the training group. They subsequently joined a two day workshop on mnemonic techniques (see 3.2.2) and came back for the second session within three days of the course finishing. One subject did not show up for the training course, and later reported having caught flu as the reason for his absence and subsequent dropping-out of the study.

After the training course, subjects continued with at-home training via an online platform (see 3.2.3). They were instructed to practise mnemonic techniques for a total of at least 24 hours over the following six weeks and received training plans with suggestions for daily training tasks to accomplish. Subjects were allowed to undertake the training during the week in accordance with their own schedule, but had to equally distribute it over the following six weeks. Subjects were invited for the third and final session after completing the training, at the earliest six weeks after the course. Compliance to the training plan was monitored via the logs of the online platform and subjects were reminded of their training when necessary:

- If a subject did not do at least four hour of training during a week, he was reminded by email to complete the training quota.
- If a subject missed the training quota for a second time in a row, he was reminded by a telephone call.
- If a subject did not finish enough training hours by week six but was not far behind, he was instructed to keep up the training and was invited for the final session when the minimal training criterion was reached.

Subjects who did not practise for at least 20 hours within a maximum of nine weeks after the training course were excluded. This applied to a total of six subjects. Thus, the training group ultimately consists of 13 subjects who successfully completed the training and underwent all three examination sessions.

Control Group

15 subjects were placed into the control group and informed that they were on the waitlist for the training and would be invited if someone cancelled. Subsequently controls were invited to the second examination session about a week after the initial session, and to the third and final examination session at the earliest six, and at the latest eight weeks after the second session. One control subject withdrew from the study after the second session stating that he would be moving abroad for an internship and not be in Munich during the time frame for the last session. Therefore the control group ultimately consists of 14 subjects who completed all three examination sessions. Participants from the control group were invited to take part in the mnemonic instruction course after completion of the final session as part of compensation.

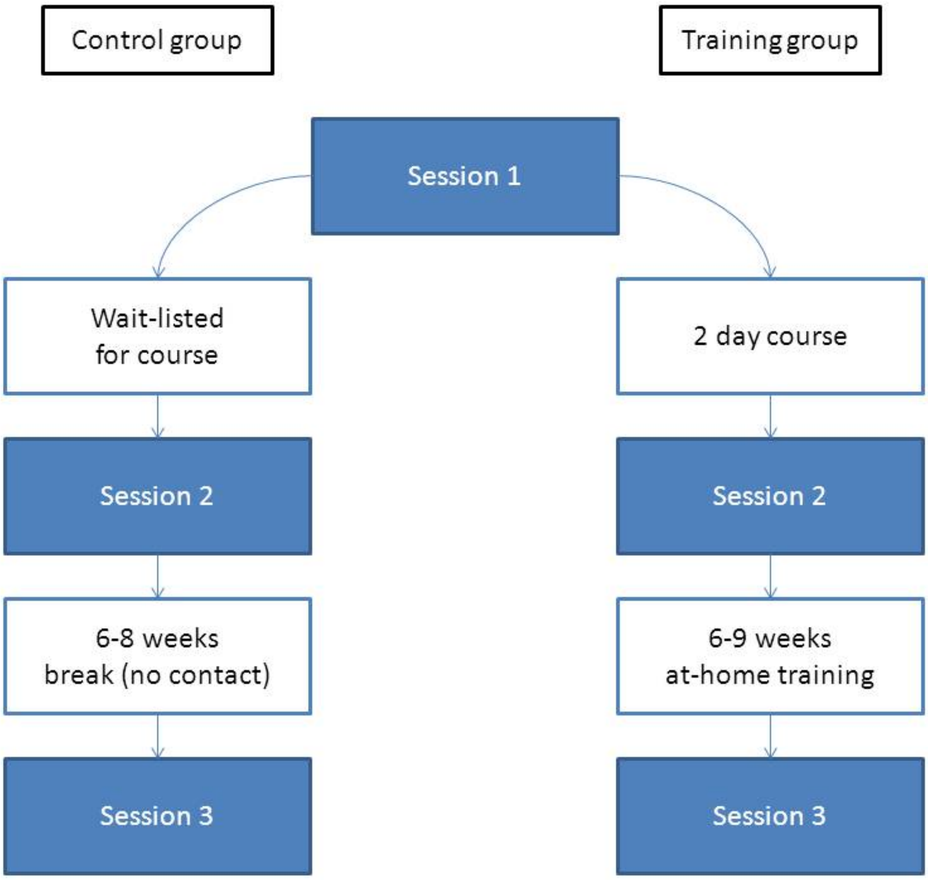


Figure 19: Study design. Subjects where either put in the wait-list control group or the training group after the first session, during which baseline measurements were taken. The second session was completed directly after the two-day instructional course in mnemonic techniques or no-contact for controls; the third session was completed after a total of at least 20 hours of mnemonic training at home, at the earliest six weeks after the course for the training group subjects or after a break of at least six weeks for the control group subjects.

3.2.2. Training Course

Subjects in the training group took part in a two-day workshop on mnemonic techniques held by the author. The maximum group size was seven participants and besides members of the training group only previous members of the control group who had finished the whole study and students from the Max Planck Institute of Psychiatry's Neuroimaging Research Group took part. The program was scheduled from 9am to 5pm on both days.

The course consisted of the following elements:

- Introduction

- Basics of Memory and Learning
- Visual Imagery
- Keyword Mnemonic
- Story Mnemonic
- Method of Loci including generation of two sets of locations, the first one with 50 locations and the second one with 25 locations
- Phonetic Mnemonic / Major System including a table of 100 images
- Memorizing Faces / Names
- Excursus: Learning Techniques, Mind Mapping, Goal Setting
- Training plan for the following six weeks
- Introduction into the at-home training platform Memocamp

All the mnemonic techniques taught were demonstrated by examples and practiced with training tasks or memory games. Comprehension and correct use of the mnemonics by all participants was tested in the seminar and additional advice and help were provided where necessary.

3.2.3. At-home training

Subjects received training plans for six weeks, consisting of instructions for the at-home training. Those included the instruction to generate further locations for use in the method of loci and training of the 100 images for the phonetic mnemonic. Participants were asked to note the amount of this offline training which they engaged in, and up to four hours were credited for the training criterion. Most of the training consisted of repeated training in the disciplines “memorizing lists of words” and “memorizing lists of digits” via the online platform Memocamp[10]. Memocamp is a commercial platform operated by a Berlin based entrepreneur. Participants received a free account and the exact amount of training including the date and time, duration, trained discipline and training success were logged with the

[10] www.memocamp.de, operated by Michael Gloschewski, Torweg 81,13591 Berlin

platforms logging function and immediately made available to the experimenters directly via the online platform.

The platform has a range of tools to support the user in training. For example, it allows the user to name the set of locations used and to track errors. This enables the user to see if a particular location or a particular image for a specific number has been forgotten or mistaken more often than others. At the beginning of the training, the user can display his mnemonic aids while memorizing when not yet familiar with the locations. The user can also set a metronome to do paced rather than self-paced memorization so to encourage faster progress. Participants of the study were instructed in the use of this platform during the training course and were asked to make use of the platforms tools while training to optimize training success. They were also asked not to use the display options anymore after week three, since they wouldn't be able to use any aids during retesting or in real-life. The participants were encouraged to contact the experimenter at any time if they encountered difficulties either in using the training tool or when training in the mnemonics.

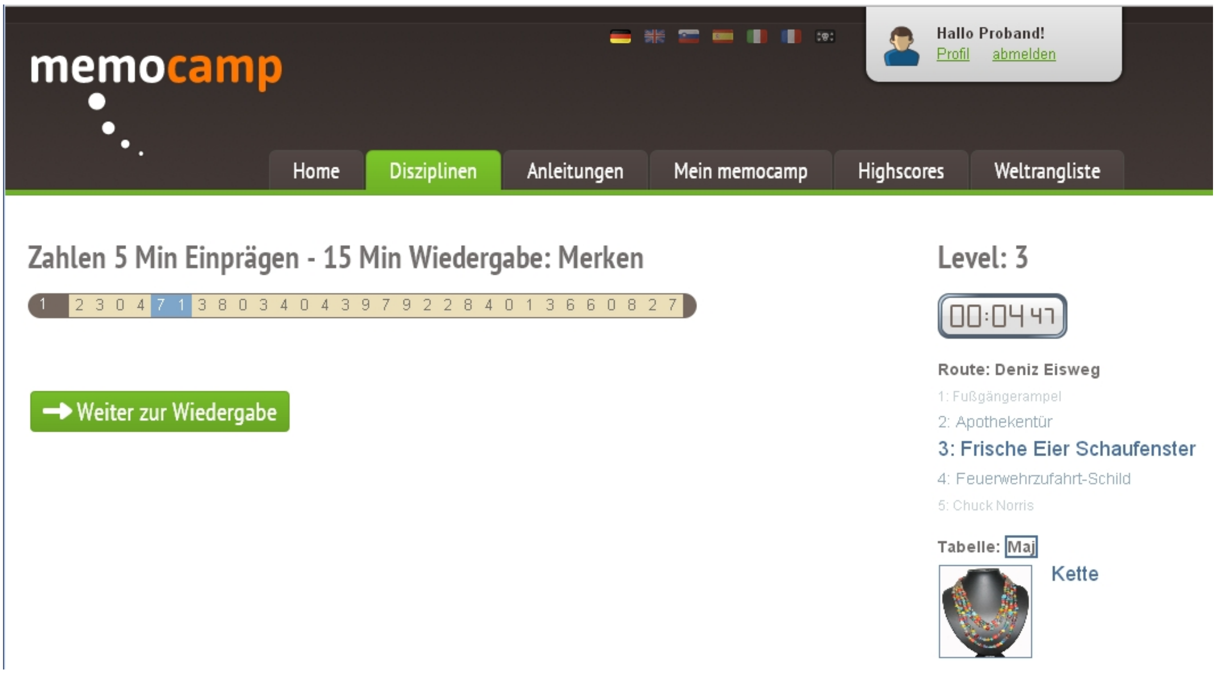


Figure 20: Screenshot of the training platform Memocamp. The participants used this platform for at-home training. It runs in the web browser and offers various tools to support practice. Displayed is a screen during a digit memorization session. The to-be-remembered string of digits is displayed in the main part of the screen with

the current two digits in focus being highlighted. The corresponding image of the Major system table for the digits, in this case 71 = Kette (German word for necklace), is displayed in the bottom right; the current location from the memory palace in use is given above it. Participants were asked to reduce the use of these aids over time while becoming more familiar with the images and locations from memory themselves.

3.2.4. Measures

		Session 1	Session 2	Session 3
Cognitive measures	Zahlen-Verbindungs-Test (ZVT)	X		X
	Bochumer Matrizentest (BOMAT)	X		X
Memory measures	Self-paced digit memorization (5 min)	X	X	X
	Self-paced words memorization (5 min)	X	X	X
	Digit span task	X	X	X
fMRI tasks	LOCI task	X	X	X
Questionnaires	Subjective memory questionnaire	X		X
	Strategy use questionnaire	X	X	X

Visual imagery questionnaire (VVIQ)	X		
Motivation type questionnaire	X		
Training questionnaire (only training group)			X

Table 3: List of the administered tasks and questionnaires.

Zahlen-Verbindungs-Test (ZVT). The ZVT (Oswald & Roth, 1987) was used as a brief measure of general cognitive ability. Tue ZVT is a trail-making task that measures mental speed and correlates highly with standard psychometric tests of intelligence. Numbers from 1 to 90 are given on a sheet of paper and have to be graphically connected in ascending order as fast as possible. The test was performed in a single-admission-version. Four trials were performed and mean scores calculated. The ZVT was done during Session 1 and Session 3.

In addition to the ZVT, the similar Trail Making Task version B (Tombaugh, 2004), was administered using parallel versions 1 and 2 randomly distributed. Due to high number of errors with just one trial per session, it was not further analyzed.

Bochumer Matrizentest (BOMAT). “BOMAT - advanced short version” was used as a measure of fluid intelligence (Hossiep, Hasella, & Turck, 2001). It is a matrix reasoning task frequently used in training studies (Jaeggi et al., 2008). The test was run in its full 45 minute version. Parallel versions A and B exist and were used with random, crossover distribution. BOMAT is aimed at assessing the fluid intelligence of above average performers and is therefore normed against university students and graduates. The norm population mean is thereby superior to the general population mean. Scores are expressed on an IQ scale with norm population mean = 100 and SD of 15. BOMAT was performed during Session 1 and Session 3.

Self-paced digit memorization (5 minutes). A time-limited and maximum power memory task for digits was performed, which closely resembled typical tasks used in the evaluation of superior memorizers. 200 digits were presented on a sheet of paper in rows of 20 digits each, with the accompanying instruction to memorize the digits in correct sequence. Participants had five minutes to memorize as many digits as possible. During recall, participants had to report the digits on a recall sheet that contained rows of 20 empty boxes. Recall time was limited to five minutes. Only digits in correct order were counted with obvious omissions ignored. The test was performed in all three sessions, employing randomly assigned differential versions.

Self-paced words memorization (5 minutes). A similar task was performed with random words. 100 German words were presented in columns of 20 words with the accompanying instruction to memorize as many words in order as possible within five minutes. Recall was performed via a recall sheet containing empty boxes; recall time was limited to five minutes. Only words recalled in the correct order were counted with obvious omissions again ignored. This test was also performed in all three sessions, employing randomly-assigned differential versions with equal difficulty in terms of word length and word frequency.

Digit Span Task (DS). A visual forward digit span task was performed on a computer screen. Digits were displayed in black font on an otherwise empty white screen, at a pace of one digit every two seconds. The test started with sequences of two digits. There were two sequences per length. After the final digit of a sequence the word “Wiedergabe” (Recall) was displayed, and recall was performed via an empty sheet of paper. Participants had to try attempt the task at least until a list with a length of 10 digits was presented, and could stop thereafter if they were sure they would not be able to complete a whole sequence of the following length, and had additionally missed at least three consecutive sequences at that time. Digits did not repeat in direct succession. The longest sequence was counted as the participant’s digit span. The test was again performed in all three sessions, and again employed randomly-assigned differential versions. The test stopped if a digit span of 15 was reached.

fMRI task – LOCI. A functional MRI scan was done while subjects performed a memory task in the scanner. Subjects could see visually presented stimuli presented on a screen by a projector. The task was programmed in "Presentation" by Neurobehavioral Systems Inc.[11] and were based on the task used in the Maguire study on memory athletes (Maguire, Valentine, et al., 2003). The paradigm consisted of a classical block design. Subjects had to memorize numbers with four digits, words and faces. After an instruction screen, subjects saw series of six numbers consisting of four digits, series of six faces that were all neutral male faces without background, including faces from the AR Face Database (Martınez & Benavente, 1998) and series of eight words (concrete nouns with three to five letters). Each item was displayed for four seconds. After each sequence, a recall block followed. Two items were presented at the same time next to each other. The subjects had to press a button to indicate, whether the left or the right one came first in the sequence just seen. Recall time for each question was five seconds. Subjects were aware that a free recall and recognition task would follow after the scanner session. After recall, a non-memory-control block followed, where the subjects saw items of the same kind as before, however this time with the instruction not to memorize but to pay attention for visual changes in the stimuli. Only two items were presented alternately and visual changes would occur only rarely (once per session per stimuli type). The paradigm began with digits followed by faces and then words. The whole design was repeated five times meaning a total of 30 four-digit-numbers, 30 faces and 40 words being presented during the task. The task was repeated in all three sessions with different versions in random crossover distribution.

[11] http://www.neurobs.com/

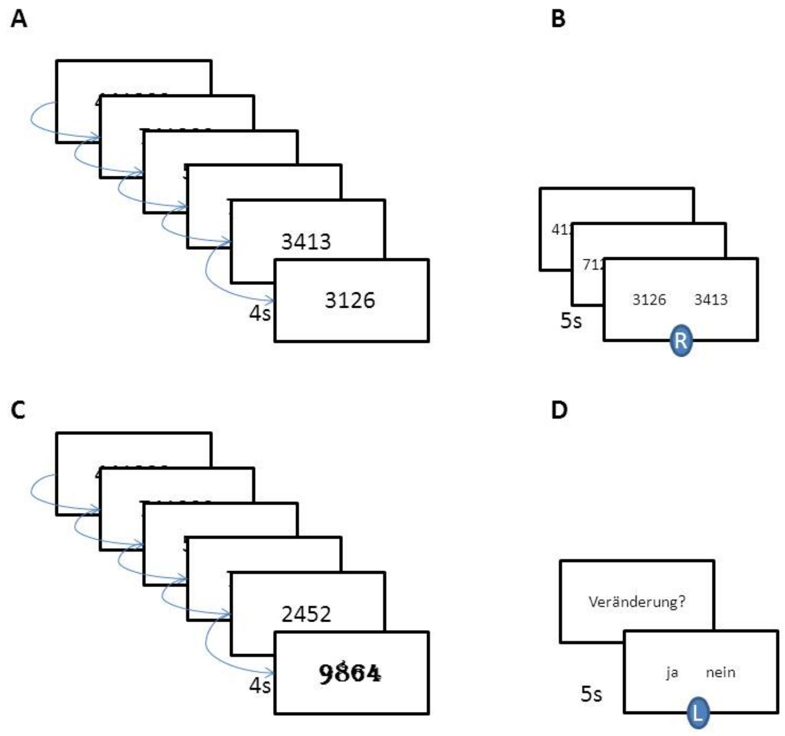


Figure 21: Scheme of one block of elements in the LOCI Task done while undergoing fMRI.

A: six elements (here four-digit numbers) were displayed after each other for four seconds each. B: All elements are recalled. Recall is done by questions on the sequence of two items displayed aside. Recall time per question was five seconds. C: control-task after the memorization. Elements of the same kind are shown, but only two in alternation. Instruction is to visually pay attention, but not to memorize the items. Sometimes the item would change optically, i.e. the font would change for words and digits or a blur effect would distort a picture. D: After the control task a question is asked, whether or not there was a change.

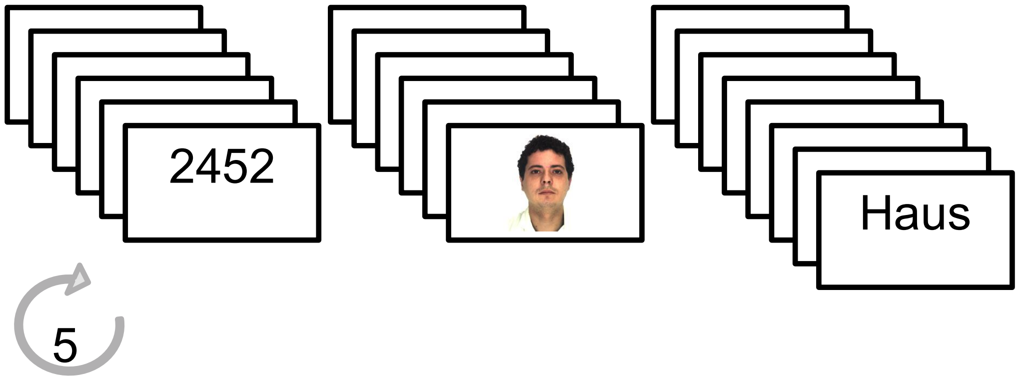


Figure 22: Schema of the whole task. The first block was number memorization (6 items), followed by face memorization (6 items) and word memorization (8 items). The whole procedure repeated five times for a total of 30 numbers, 30 faces and 40 words.

After the scan an empty sheet was handed out for free recall of remembered four digit numbers and words from the scan. Only completely correct items from the memorization blocks were counted. Afterwards recognition sheets were handed out with all presented items (numbers, words, faces) and the same number of distractors being shown. For each item the subject had to answer if the item was presented during the scan with “yes” or “no” and how sure they were with their judgment with the four possibilities “I was sure”, “I am rather sure”, “I am rather unsure”, “I guessed”.

Questionnaires. At the beginning of the first and third session a self-assessment questionnaire was handed out. Subjects were asked to rate a) their general memory ability, b) their memory ability for digits, c) their memory ability for names and faces and d) their memory ability for written information by placing a mark on semantic differential scale with the extremes “schlecht” (bad) on the left and “gut” (good) on the right scored from 1 to 10. Subjects were also asked to estimate how many digits and how many words they will be able to memorize in correct sequence within five minutes.

At the end of all three sessions, a strategy use questionnaire was provided asking to briefly describe the strategy use for all memory tasks done.

A self-translated German version of the "Vividness of Visual Imagery Questionnaire" (VVIQ; McKelvie, 1995) and a motivational questionnaire based on the self-translated twelve questions of the Achievement Goal Questionnaire (AGQ; Elliot & McGregor, 2001) plus three expectation questions on expected success in the memory training and two questions on motivation for joining the study (money and memory training) on joining the study using a seven point Likert scale were administered at the beginning of Session 1. A questionnaire on the estimation of how many hour spent training, an open question on what kind of training they did outside of the Memocamp training platform and an assessment of how much they enjoyed training was handed out to the training group participants only in session 3.

3.2.5. Data analysis

Analysis of behavioral was done using SPSS 18. Data is reported as mean ± standard deviation (s.d.) despite where otherwise noted. Significance was assumed for an alpha of 5% and is reported in steps of * = $p < .05$, ** = $p < .005$ and *** = $p < .001$.

A one-way ANOVA was done do compare the groups (trainings, control, drop-outs) at pre-test to look for group differences. Repeated measures ANOVA with the between-subjects factors group (training, control) and within-subjects factor time (TP1 pre training, TP2 after instruction TP3 post training; no contact for controls) were done for the various training and transfer tasks. Post hoc t-tests were done were significant main effects were found to look into differences between groups. Two-tailed t-tests where applied were no specific assumption of a training effect was given and one-tailed t-tests were used where such could be expected. Effect sizes are either given as η_p^2 or Cohen´s d.

For ANOVAs homogeneity of variances was assessed by Levene's Test of Homogeneity and homogeneity of covariances via Box's test of equality of covariance matrices. The assumption of sphericity was tested via Mauchly's Test for Sphericity and Huynh-Feldt correction was used when violated.

Where correlations between variables where investigated, Pearson correlations were used despite when normality was violated. In this case non-parametric Spearman correlation was done. When multiple correlations where tested (e.g. when comparing memory task improvements with transfer task improvement), Bonferroni correction was done to correct the significance level.

3.2.6. fMRI data acquisition and analysis

fMRI was carried out at 3 T (Discovery MR750, GE Healthcare, Waukesha, WI,USA) using an 12-channel head coil and covering 42 AC-PC oriented slices (2 mm thickness, 0.5 mm gap; 128 × 128 matrix, interleaved echo planar images, TR 2500 ms, TE 30 ms). fMRI analysis was done with Matlab2008b[12] and SPM8 software[13].

Preprocessing

The function images were preprocessed using SPM8 and consisted of the following steps: (1) correction of slice time differences due to interleaved images acquisition, (2) realignment to the first volume using rigid body transformation, (3) normalization to the EPI template in Montreal Neurological Institute (MNI) space, (4) resliced (voxel resolution 2 × 2 × 2 $mm^{3),}$ smoothed using a Gaussian kernel of 6 mm full-width at half maximum (FWHM) and (5) segmentation in native space. The first four images were discarded after preprocessing to remove non-steady-state effects.

Analysis

A full factorial model was calculated using the factors time (TP1, TP2, TP3), condition (numbers, faces, words) and group (training, controls). The three-way time x condition x group interaction was calculated at cluster level FWE correction with a threshold of $p < .05$; sampled at $p < .001$ uncorrected and followed-up by direct group comparisons (t-tests) applying the same thresholds.

[12] By MathWorks, http://www.mathworks.de/products/matlab/
[6] http://www.fil.ion.ucl.ac.uk/spm

3.3. Results

3.3.1. Subjects & Group differences

	Age (years)		BOMAT (IQ sc.)	
	mean ± s.d.	range	mean ± s.d.	range
Training Group (n = 13)	22.9 ± 3.3	19 – 28	101.8 ± 12.7	88 - 118
Control Group (n = 14)	22.9 ± 2.2	19 - 26	102.0 ± 14.1	78 - 122
Dropouts (n = 7)	24.4 ± 3.4	20 – 29	101.1 ± 10.5	85 - 114

Table 4: Statistical data on the subjects randomly split into a training group and a control group. Dropouts were those who either did not come to the training (one subject) or did not fulfill the training plan within eight weeks (six subjects). BOMAT is given in IQ scale (mean 100, s.d. 15) but normed on students and graduates only, not on the general population, so not representing the IQ itself.

Statistical data on the subjects is given in Table 4. A one-way ANOVA was conducted to compare the groups with no statistically significant differences to be found for age ($F(3,31) = .792$, $p = .462$) or fluid intelligence ($F(3,31) = .010$, $p = .990$) between the training group, control group and dropouts, who did not finish the training.

All subjects completed the Achievement Goal Questionnaire with added questions on expected success and enjoyment in the training as well as questions asking for the role of the monetary reward (only paid when finishing the whole study) and the role of the memory training offered for their decision to join the study. Another one-way ANOVA was conducted to compare groups on their answers to these questions. There was a significant main effect of group for influence of monetary reward ($F(2,30) = 4.567$; $p < .05$). Post-hoc t-test (two-tailed) revealed that drop-outs where significantly less interested in the money ($p < .05$) than those who finished the training fitting to their decision to abort the study and by that relinquish the money. There was no difference between the training and control group ($p = .60$). There was neither a difference between the groups in valuation of the memory training ($F < 1$) nor expectancy of success in the training ($F < 1$) nor any motivational domain.

The training of the training group subjects was monitored using the web-based training platform Memocamp. Subjects were asked to report additional training (including generating new locations for the method of loci and training the images for the phonetic mnemonic) outside of Memocamp in hours. A minimum of 20 hours of training within eight weeks was necessary to be invited to the post-test (latest nine weeks after the training course). Based on Memocamp and outside training, in total the training subjects had trained using mnemonics for 24.22 ± 3.27 hours.

Memocamp did not log training hours for one of the disciplines used, hence this value is slightly underestimating the real total training time.

3.3.2. Training improvement in memory tasks

Analysis of the results in the self-paced number memorization task (5 minutes), self-paced word memorization task (5 minutes) and digit-span task show that the training group achieved marked improvements (see Figure 23).

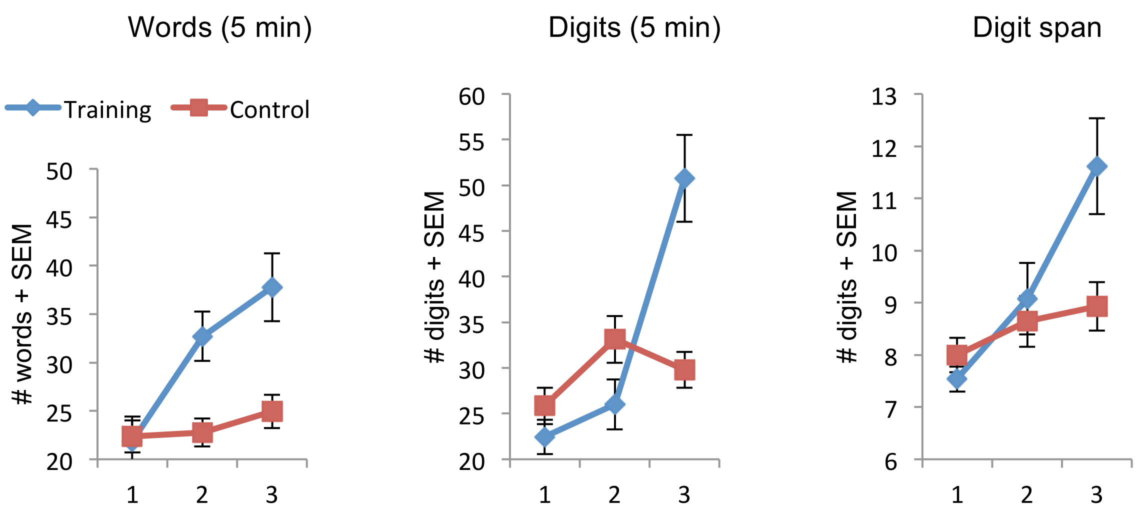


Figure 23: Memory task performance (mean performance and standard error of the mean) by training group (blue) and control group (red) at time points (TP1) before training, (TP2) after introduction weekend / waitlist and (TP3) six to nine weeks of training / break.

Data was analyzed with repeated measures ANOVAs[14] with the factors time (TP1 pre instruction, TP2 post instruction and TP3 post training) and group (training and control) for each task which revealed a main effect of group only for words ($F(1,25) = 6{,}788$; $p < 0.05$; $\eta_p^2 = .214$) and main effects of time for all three tasks. Much stronger improvements in the training group were confirmed by highly significant time * group interactions for all three tasks with the strongest interaction in digits ($F(2,50) = 16.777$; $p < 0.001$, $\eta_p^2 = .402$) followed by words ($F(2,50) = 13.031$; $p < 0.001$; $\eta_p^2 = .343$) and digit span ($F(2,50) = 8.854$; $p = 0.001$; $\eta_p^2 = .262$). Post hoc ANOVAs for each group showed that there was a significant time effect only for digits in the control group ($F(2,26) = 4.439$; $p < 0.05$; $\eta_p^2 = .255$) but for all three tasks in the training group. Performance improvement and effect sizes for the training group are given in Table 5.

[14] Testing for outliers via boxplot inspection showed two outliers in the digit task at time point two within the training group and one outlier in the word task at time point three within the training group (values greater than 1.5 box-lengths from the edge of the box), but none of them was extreme (above 3 box-lengths). Therefore the outliers were kept within the analysis. Shapiro-Wilk testing for normality showed that all data was normally distributed (p > .05) with one exception for digit span at time point one within the training group (p = .025). Since a ANOVA is fairly robust to slight deviations from normality (Schmider, Ziegler, Danay, Beyer, & Bühner, 2010), further analysis was done regardless.

Task	Time Point	Mean	s.d.	Cohen´s d
Digits	1	22.46	6.83	
	2	26.00	9.90	.4
	3	50.77	17.15	2.2
Words	1	21.85	9.15	
	2	32.69	9.19	1.2
	3	37.77	12.66	1.4
Digit Span	1	7.54	.87	
	2	9.08	2.47	.8
	3	11.54	3.231	1.7

Table 5: Results in the memory tasks in the training group and effect sizes for the improvements compared to time point 1 (before instruction).

Post hoc t-tests (two tailed) revealed significant differences between the groups in the digits task at TP3 ($t(25)=4.189$; $p < 0.001$; Cohen´s $d = 1.6$), in the words task at TP2 ($t(25)=3.438$; $p < 0.005$; Cohen´s $d = 1.3$) and TP3 ($t(25)=3.356$; $p < 0.005$; Cohen´s $d = 1.3$) and in the digit span at TP3 ($t(25)=2.954$; $p < 0.01$; Cohen´s $d = 1.1$) with very large effect sizes. Digit span group differences and training gains might have been bigger: Five out of 13 subjects of the training group reached the ceiling digit span of 15 digits (14 in one case due to software failure) possible in the test administered at TP3 after training whereas no control subject did (best control subject reached a digit span of 12). No subject reached this limit at TP1 or TP2.

Top three performances in the digit memorization task were 83, 74 and 70 digits, all at the third time point by subjects from the training group. Top three performances in the word memorization task were 70, 51, and 45 words, again all by training subjects at the final time point.

To test if initial performance influenced training gains, two-tailed Pearson correlations were done for performance at pre-test and training gain in each task separate for both groups. A non-parametric two-tailed Spearman correlation was done for digit span initial performance and digit span training gain due to violation of the assumption of normality in this task in the training group. For the Control group a strong negative correlation was found between initial performance in the digits (5min) task and the improvement in this task ($r = -.686$; $p < .001$) indicating a compensation effect, i.e. the better subjects did not improve as much as those with poorer initial scores. There were also slight negative, but not significant, correlations for words ($r = -.340$; $p_{two\_tailed} = .234$) and digit span ($r = -.359$; $p_{two\_tailed} = .208$) in the control group. In the training group no significant correlations were found ($p_{digits} > .3$, $p_{words} > .7$, $p_{digitspan} > .7$) indicating that the training gains by mnemonic training were independent of initial ability level and no compensation effect being present (see Figure 24).

Correlations were also calculated between training gains in the different tasks and pre-test ZVT score and pre-test BOMAT score, but were not significant when corrected for multiple testing ($r < .3$; $p > .4$ for words and digits with ZVT and BOMAT correlations), but there was a trend for pre-test ZVT and digit span improvement at $r = -.630$ and $p_{uncorrected} = .021$. Further, also a median split on BOMAT performance at pretest within the training group was done to assess if intelligence was driving differences. Lower performer range was 11 to 15 correct BOMAT items, higher performer range was 19 to 21 correct items equaling lower IQ scale range of 85 to 100 and higher performer IQ scale range of 112 to 118; remember that BOMAT is not normed against population but students and graduates. ANOVAs for each measure (words, digits, digit span) showed no performance group by time interaction (words $F(1,11) = .064$; $p = .805$; digits $F(1,11) = .873$; $p = .370$; digit span $F(1,11) = 2.874$; $p = .118$) indicating that memory improvement by mnemonic strategy training was independent of intelligence.

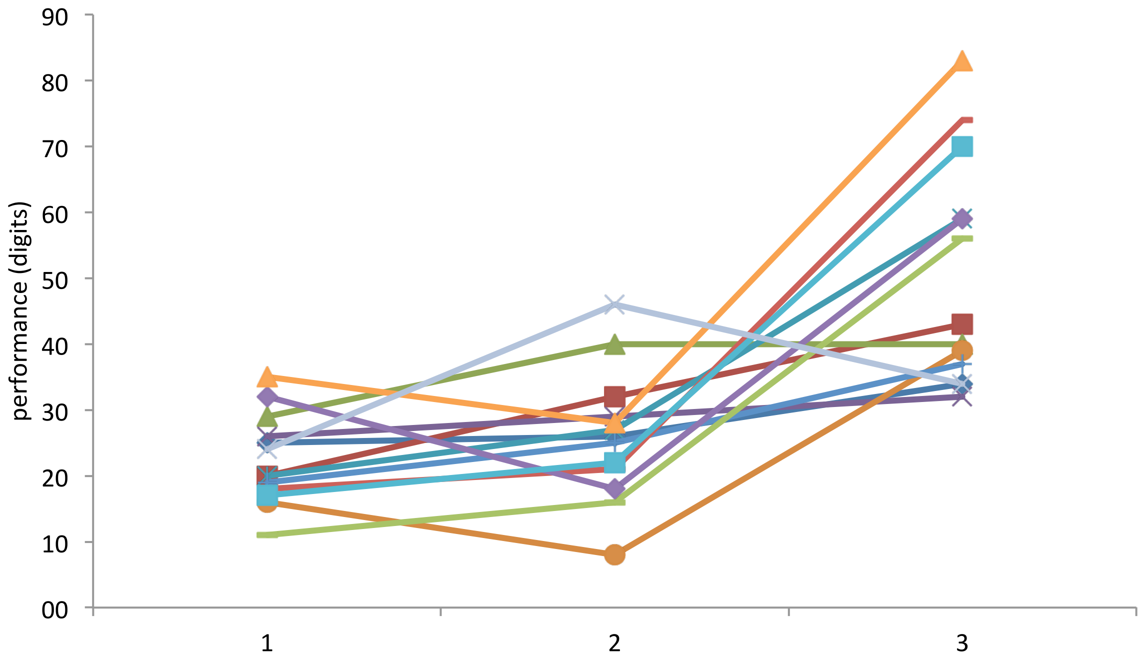


Figure 24: Performance in the digit span task (five minute self-paces memorization; free recall afterwards) for all 13 subjects of the training group at the three time points. Training gains did not correlate with initial performance but all subjects improved comparably regardless of initial performance.

Training data of the subjects was collected within Memocamp, but options used (like displaying images and loci as aids, using metronome function) were not logged and individuals differed in using these options. Thus training data could not be assessed group wise. Visual inspection showed rather linear improvements in most subjects. An exemplary training curve of the most successful subject in the digit memorization task is given in Figure 25. The best trial per training day is given. The subject achieved a personal best of 190 digits correctly remembered after five minute of memorization time in the online tool (note that the training tool displays rows of 40 digits thereby motivating to attempt full rows probably explaining the small plateau before reaching 80).

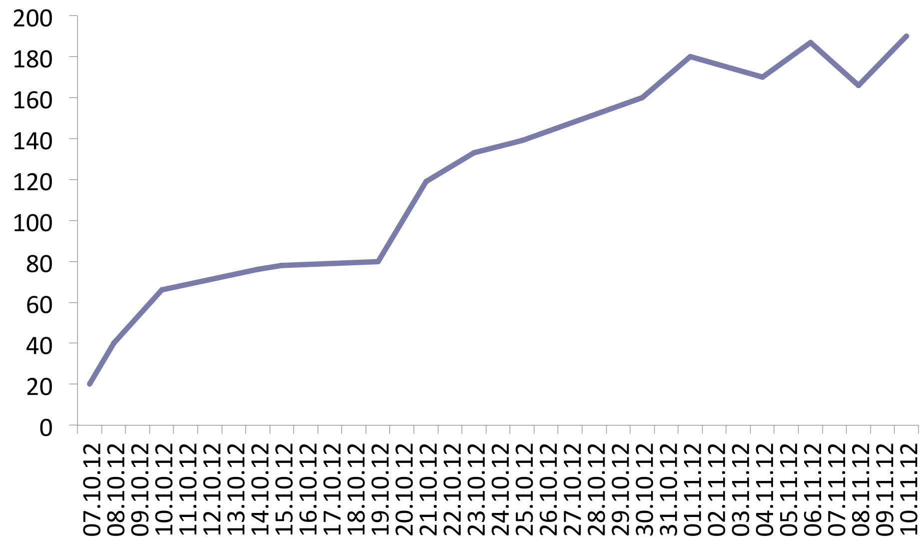


Figure 25: One subject´s exemplary at-home online training curve (best score of each day in the digit memorization task).

3.3.3. Behavioral improvement in LOCI task

Order Recall

Answers given by button presses in the scanner were not fully logged due to technical malfunction twice, once for a training subject and once for a control subject. These two were excluded from the analysis in this section only.

In the order recall done within the scanner, repeated measures ANOVAs with the between subject factor group and within subject factor time revealed significant group x time interaction for words ($F(2,44) = 6.022$; $p < .01$; $\eta_p^2 = .215$) but not for numbers or faces ($F(2,44) < .5$ in both). Main effects for time or group were not significant either (main effect of time for faces $F(2,44) = 2.090$; $p = .136$; main effects of time for digits and words $F(2,44) <$ 1.0, all three main effects of group $F(2,44) < 1$).

Post hoc t-test revealed the group difference between the training and the control subjects for words at TP3 to be significant ($t(23) = 2.117$; $p < .05$).

Free Recall

In the free recall condition of the LOCI task done after the scan, training subjects showed a strong improvement in the words condition but not in the numbers condition (see Figure 26), where both groups showed floor effects with many subjects (15 at TP1, 10 at TP2 and 8 at TP3) unable to correctly recall a single four-digit number. In contrast, ceiling effects played a role in the words condition. Five subjects at TP2 and six subjects at time TP3 scored 35 or more out of 40 possible words.

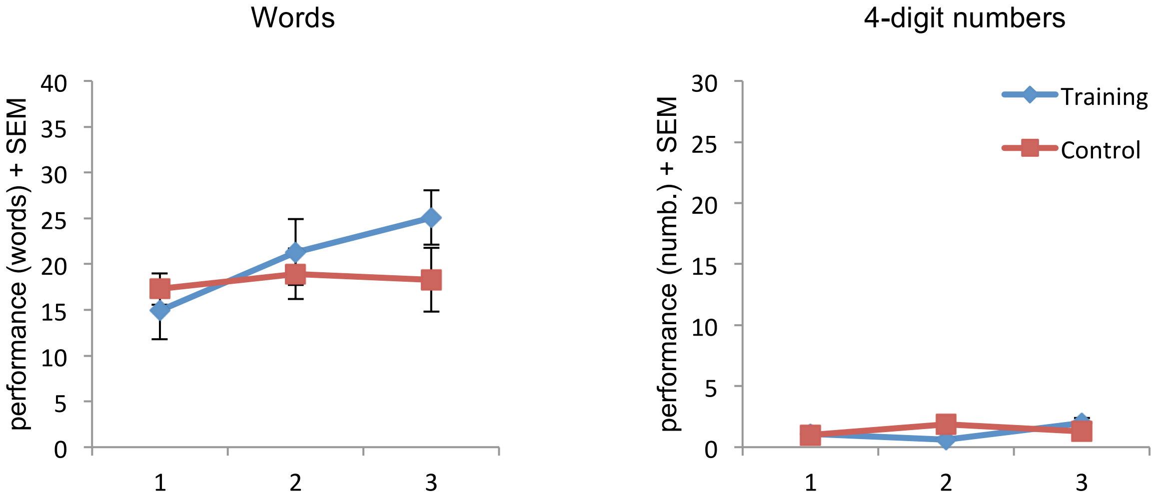


Figure 26: Mean performance (and standard error of the mean) in the written free recall after encoding in the scanner at time points (1) before training, (2) after introduction weekend / waitlist and (3) six to nine weeks of training / break.

Due to the strong floor effects, the performances in the free recall of the numbers, performance data for this condition violated the normality assumption as assessed by Shapiro-Wilk test ($p < .05$ in both groups at all-time points with the exception of training group at TP3 where $p = .068$). Since a transformation of data was unsuccessful, only visual inspection of the graph (see Figure 26) was done, suggesting no group difference.

Box's Test of Equality of Covariance Matrices shows a violation of homogeneity of covariances ($p < .05$) and separate repeated measures ANOVA were run for both groups which revealed a significant time effect for the training group ($F(2,24) = 4.422$; $p < .05$; η_p^2 =

.269) but not for the control group (Huynh-Feldt correction for violation of sphericity, $\varepsilon = .77$; $F(1.5,19.9) = .228$; $p = .740$). Post hoc t-tests (one-tailed) revealed a significant improvement for the training group from TP1 before instruction to TP3 after several weeks of training ($t(12) = 3.3$; $p < .005$) with a strong effect size (Cohen´s d = .9). The improvement from TP1 to TP2 was also significant ($t(12)=2.1$;$p < .05$, Cohen´s d = .5), however the further improvement from TP2 to TP3 was not ($p > .05$), maybe due to the ceiling effects at TP3.

Recognition

After free recall, participants had to fill out recognition sheets for all three conditions with 60 items presented for numbers and faces (the 30 stimuli seen during the task plus 30 distractors) and 80 items for words (40 stimuli, 40 distractors) and results are given in Table 6. Faces recognition data is missing for one subject from the control group due to non-compliance to instruction.

	Time Point	**Training**		**Control**	
		Mean	**SD**	**Mean**	**SD**
Numbers	1	57.6%	4.8%	59.9%	5.9%
	2	57.8%	6.9%	56.8%	3.9%
	3	66.3%	10.7%	58.9%	5.7%
Faces	1	72.4%	6.3%	76.1%	24.2%
	2	79.4%	12.9%	75.2%	24.6%
	3	77.8%	12.8%	72.6%	25.8%

Words	1	86.3%	11.2%	91.4%	7.5%
	2	89.3%	11.8%	92.2%	8.7%
	3	91.3%	8.2%	88.1%	11.1%

Table 6: Performance in the recognition task after the scanner session. All stimuli from the scan plus the same amount of distractors were given. Subjects had to mark, if they had seen the item or not. 50% equals chance level performance.

Box's Test of Equality of Covariance Matrices shows a violation of homogeneity of covariances ($p < .01$) and two separate repeated measures ANOVA were run. One used the factors time, group and stimuli (numbers, words) while the other was concerned with faces only. Digits and words were the two most-trained disciplines during the six to eight weeks of training.

ANOVA revealed a significant main effect of stimuli ($F(1,23) = 315.2$; $p < .001$; $\eta_p^2 = .932$; with words better recognized than numbers), a trend for time ($F(2,46) = 2.45$; $p = .089$; $\eta_p^2 = .100$), a significant time by group interaction ($F(2,46) = 4.838$; $p < .05$; $\eta_p^2 = .174$), and a significant time by stimuli interaction ($F(2,46)=4.119$; $p < .05$; $\eta_p^2 = .152$).

Post-hoc t-tests (one-tailed) revealed a group difference for numbers at TP3 post training ($t(23)=1.97$; $p < .05$; Cohen´s d = .80), significant improvements in the training group from TP1 to TP3 in numbers ($t(12)=2.76$; $p<.01$; Cohen´s d = 1.05), from TP2 to TP3 in numbers ($t(12)=3.03$; $p<.01$; Cohen´s d= .94) and from TP1 to TP3 in words ($t(12)=1.93$; $p < .05$; Cohen´s d = .5). For the controls there was a significant decrease in recalled words from TP2 to TP3 ($t_{two\text{-}tailed}(12)=-2.82$; $p < .05$; Cohen´s d = .4) but no other comparison was significant in the controls.

For the faces stimuli there was no statistically significant interaction between the intervention and time (F(2,46)=1.854; p = .17; η_p^2 = .75); main effects of time (p > .5) and group (p > .3) were also non-significant.

3.3.4. Transfer effect to processing speed

ZVT scores were transformed to IQ scales corresponding to the norms for single administration (Oswald & Roth, 1987). Scores differentiate between the age group 16-20 and 21-30, and were taken on an individual basis for each subject in accordance to his age. ZVT was administered at TP1 and TP3 only.

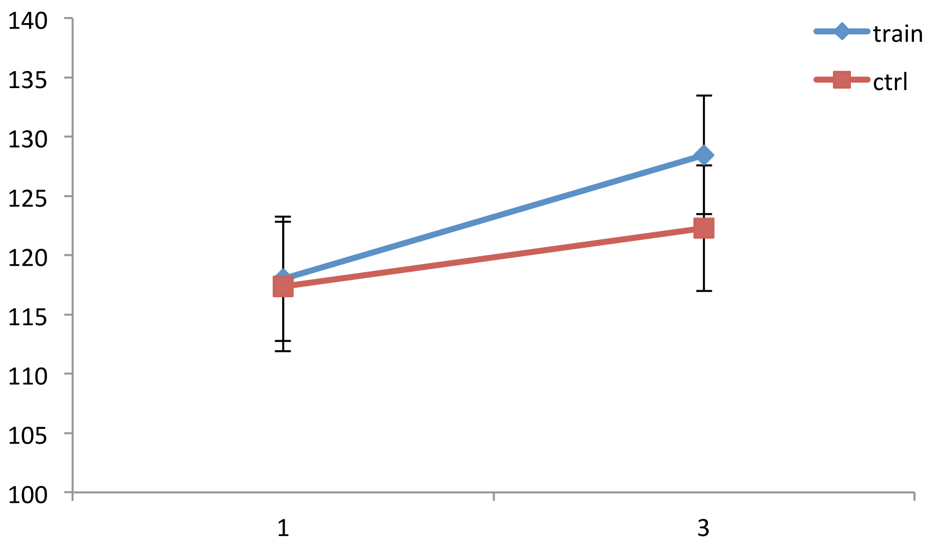


Figure 27: Training gains in the ZVT from pre instruction (1) to post training (3) as given on an IQ scale.

There was a significant re-test effect in the control group (t(13)=4.762; p < .001; Cohen´s d = .25), however a significantly higher improvement in the training group (t(12)=7.359; p < 0.001; Cohen´s d = .56) as confirmed by a significant time by group interaction (F(1,25) = 10.112; p < 0.005; η_p^2 = .288). A median split on performance at pretest within the training group was done to assess whether initial performance in the ZVT was driving differences, but that was clearly not the case (F(1,10) = .275) as both low and high performers benefited equally. Also, the improvement did neither depend on fluid intelligence as measured with BOMAT at pretest (F(1,10) = 2.414; p = .151) nor working memory capacity as measures with digit span at pretest ($F(1,10) = .550$; $p = .476$) showing that all subjects processing speed improvement due to mnemonic training was independent of their initial abilities.

When looking for correlations with performance, ZVT in seconds (faster equals better) at pretest correlated (Pearson correlations, one-tailed, Bonferroni corrected for multiple testing) with the performance in all three memory tasks ($r_{words} = -.432$; $_{pone\text{-}tailed,corrected} < .05$; $r_{digits} = -.610$; $p_{one\text{-}tailed,corrected} < .01$; $r_{digitspan} = -.501$; $p_{one\text{-}tailed,corrected} < .01$) over all subjects. At post-test this was true for words $r_{words} = -.423$; $_{pone\text{-}tailed,corrected} < .05$) and digit-span ($r_{digitspan} = -.628$; $p_{one\text{-}tailed,corrected} < .001$), but not for digits ($r=-.102$; $p > .05$), regardless if comparisons were done over the whole sample or only within the training group (r-values given for whole sample).

When looking for correlations between improvements in the ZVT with improvements in the three memory tasks within the training group, none was found to be significant ($p > .3$ for all three).

3.3.5. No transfer effect to fluid intelligence

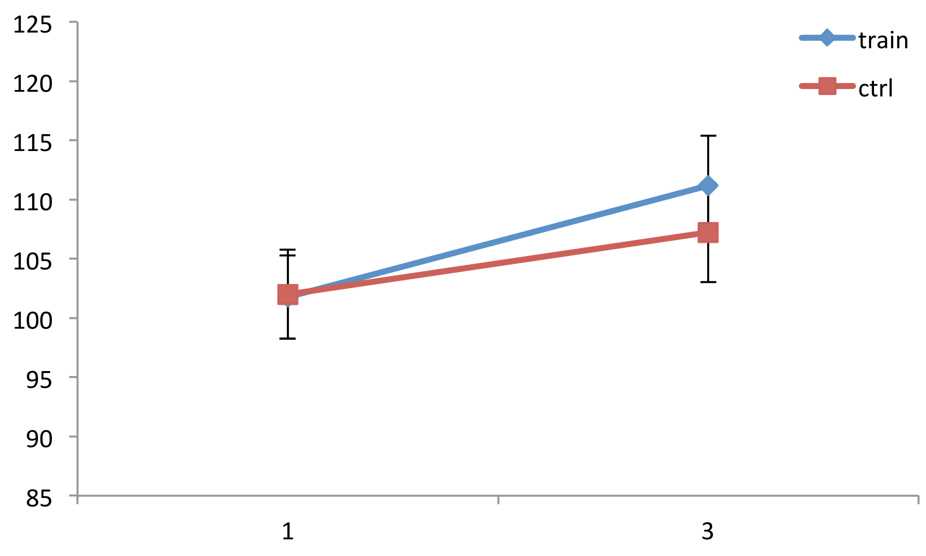


Figure 28: Training gains in the BOMAT fluid intelligence task from pre instruction (1) to post training (3) as given on an IQ scale compared against a norm population of students and graduates.

BOMAT scores on IQ scale (norm population mean 100, SD 15) were taken corresponding to the norms for single administration in 16 to 30 year old (Hossiep et al., 2001). The control group improved from 102.0 ± 14.1 to 107.2 ± 15.6 and the training group from 101.8 ± 12.7 to 111.2 ± 15.2. There was a significant main effect of time (F(1,25)=10.889; p < .005; η_p^2 = .304), but the group by time interaction was not significant (F(1,25)=.889; p=.355; η_p^2 = .034)

3.3.6. Self-appreciation of success

Training group subjects gave a subjective self-evaluation of their memory at TP1 and TP3 before doing the memory tasks. At TP1 the average estimation was to be able to memorize 15.7 ± 6.8 digits and 12.8 ± 3.7 words in five minutes. At TP3 this rose to 64.2 ± 44.8 digits and 38.5 ± 15.7 words. Compared to actual performance (see Table 7) this is a slight underestimation of both tasks at TP1, a very accurate estimation for words at TP3 and a slight overestimation for digits at TP3.

		mean	**s.d**		**mean**	**s.d**
PRE	digits estimate	15.69	6.81	digits achieved	22.46	6.83
	words estimate	12.77	3.65	words achieved	21.85	9.15
POST	digits estimate	64.15	44.82	digits achieved	50.77	17.15
	words estimate	38.46	15.73	words achieved	37.77	12.66

Table 7: Self-estimation and actual performance in the self-paced digits and words memorization tasks by training group subjects.

Despite the highly significant (digits: t(12) = 4.282; p = .001; words: t(12) = 5.802; p < .001) rise in their estimation of the number of digits or words which they would be able to memorize, participants' self-evaluation on the semantic differential scale represented only a slight increase. This was significant for the evaluation of digit memory ability ($t(12) = 2.292$; $p < .05$) and names and faces memory ability ($t(12) = 2.213$; $p < .05$) but not for general memory ability ($t(12) = 1.389$, $p = .190$) or memory ability for textually presented material ($t(12) = .634$; $p = .538$) (see Table 8)..

	pre		post	
	mean	**s.d.**	**mean**	**s.d**
General	6.0	1.1	6.7	1.4
Digits	5.4	2.0	6.7	2.0
Names and faces	5.3	1.8	6.7	1.6
Textual information	6.5	1.5	6.8	1.5

Table 8: Self-evaluation of one's memory ability in four domains on an on semantic differential scale with the ends bad (1) to good (10).

3.3.7. fMRI

Interaction and main effects

First the three-way interaction group x day x condition was calculated. Somewhat surprisingly only clusters in the cerebellum (R: 50;-72,-28; 6, -64, -20; L: -2;-76,-34) were significant at cluster level when multiple test correction (FWE p < .05; sampled at p < .001 uncorrected) was applied. When looking for a main effect of group, no significant cluster was found.

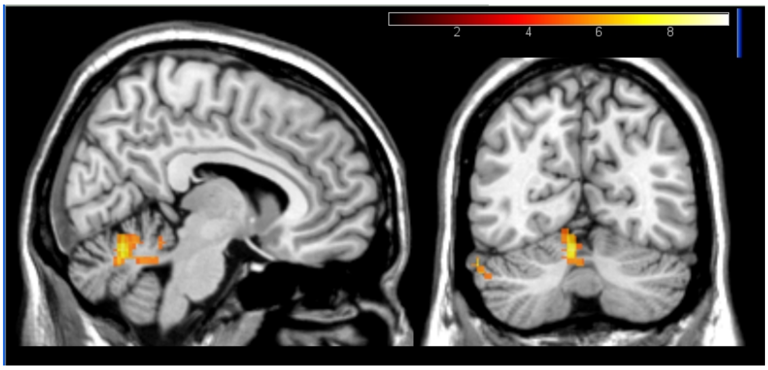


Figure 29: Three-way interaction group x day x condition at a threshold of uncorrected p < 0.001 per voxel followed by cluster based multiple test correction procedure (FWE, p < .05).

The performance in the numbers condition (both groups performed poorly with a significant difference only in recognition at TP3) and faces (both groups performed equally well) was similar between groups, whereas the training effect in words was much stronger. In the recall condition performed during fMRI the training subjects only improved in the words task (compare 3.3.3). Therefore following analysis concentrates on the words task.

The average effect of condition over both groups and all days in the words task is given in Figure 30 and Table 9.

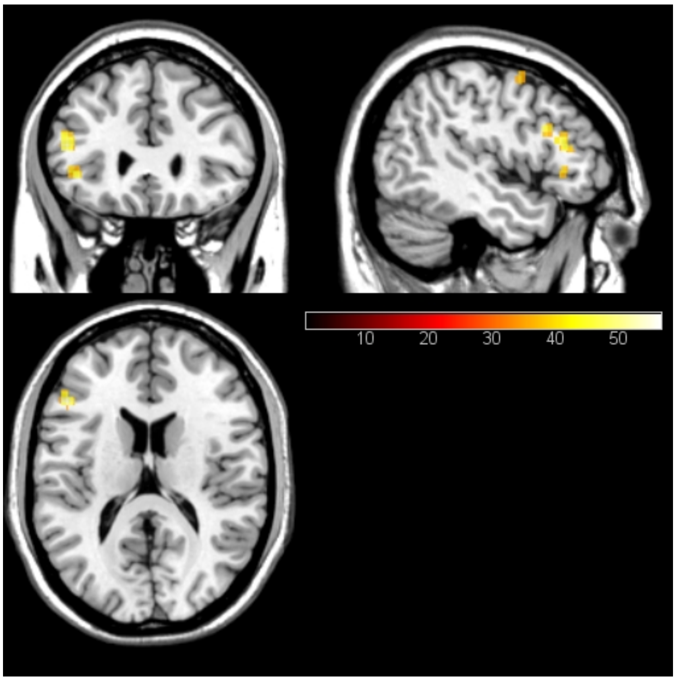


Figure 30: Average effect of condition for the words task at all three time points; sampled at voxel wise corrected $p_{FWE} < 0.05$ and an extend threshold of 15 voxels. Activations are found, beyond others (see Table 9), in Broca Area (-48, 28, 16) known for language comprehension and left middle frontal gyrus (-30, 0, 50, Brodmann area 6) associated with verbal memory.

Brain Region	Peak voxel coordinates (MNI)	Cluster size k
Middle Frontal Gyrus (L) (BA6)	-30, 0, 50	298
Broca Area / Inferior Frontal Gyrus (L)	-48, 28, 16	150
Inferior Frontal Gyrus (L)	-42, 26, -2	71
Cingulate Gyrus (L) / Medial Frontal Gyrus (L)	-6, 16, 44	64
Right Cerebellum	38, -68, -34	42
Middle Occipital Gyrus	44, -80, 6	18

Table 9: Brain regions found in average effect of condition for memorizing words over all study days.

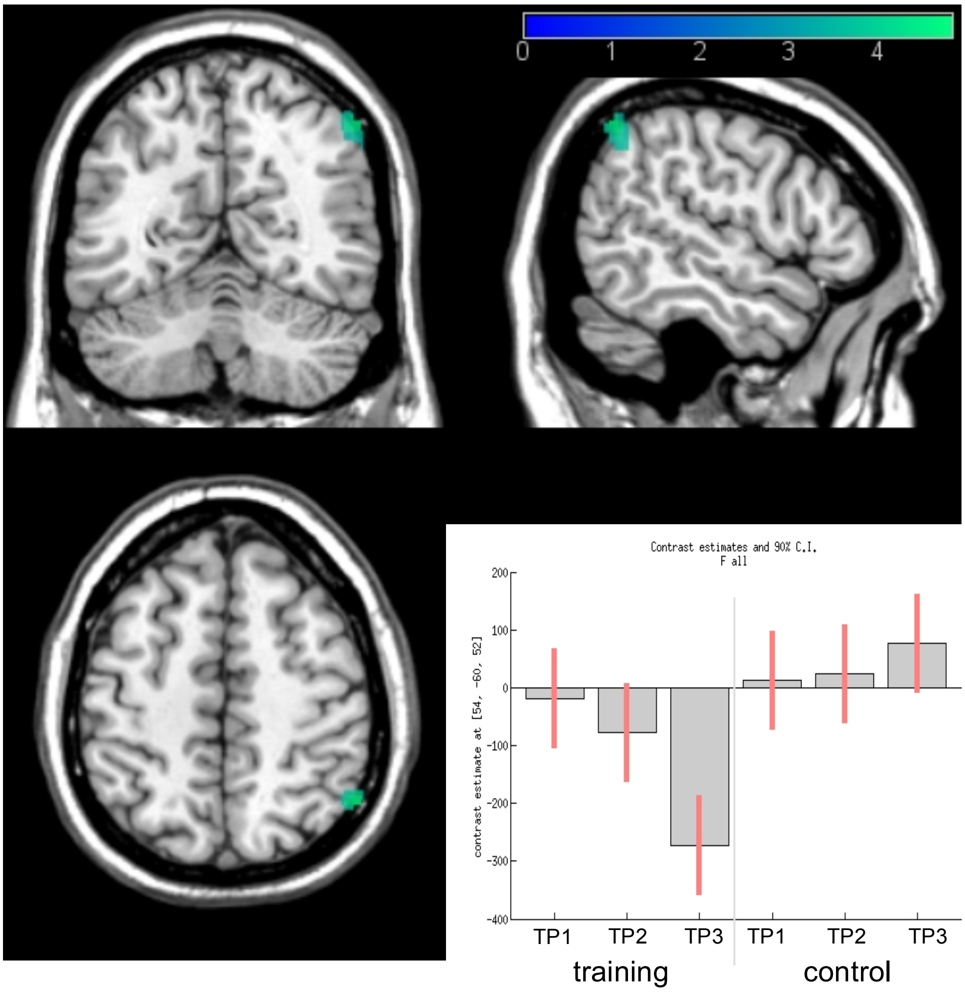


Figure 31: Activation differences between training subjects and controls at TP3 sampled at uncorrected p < .005 and FWE correction on the cluster level with p < .05. Only a significant deactivation in training subjects was found in right inferior parietal lobule / BA 40 (54, -60, 52).

3.4. Discussion

3.4.1. Memory improvement and preconditions

The study consisted of a combined training in the method of loci (see 1.3.7) and the phonetic mnemonic (see 1.3.6). During the two-day instruction course other mnemonics were explained and tested, but not further trained during the following weeks.

Method of Loci

Memory tasks used to measure memory improvement were self-paced memorization of digits and words. The method of loci has repeatedly been shown to enhance word list learning even with little training (e.g. Bower, 1970b; Massen & Vaterrodt-Plünnecke, 2006; Moe & De Beni, 2005b; Roediger, 1980). While memory athletes who show extraordinary memory performance credit their skills to the method of loci (Maguire, Valentine, et al., 2003; Study 1 of this thesis), very few attempts have been done to observe the outcome of prolonged training using this method in naïve subjects. In one study four young subjects of above average intelligence improved their memory for serial order words recall at a presentation time of 10 seconds per word from about 5 words to about 39 words after 26 training sessions of about 90 minutes, but list length was limited at 40 words. A group of 18 subjects reported in the same paper improved to 23 words out of 30 words at the same pace after 20 sessions (Kliegl et al., 1989). In the present study a group of 13 subjects achieved a mean performance of about 38 words memorized within 5 minutes (self-paced; average about 8s per word).

Phonetic Mnemonic

The phonetic mnemonic has been shown to enhance memory for digits in group studies as well (Higbee, 1997). One single case is reported (Kliegl et al., 1987), where the subject combined the phonetic mnemonic and the method of loci and achieved to memorize 80 digits at a pace of 5s/digit but could not keep up his skill at a rate of 2s/digit. Other studies hypothesized that only subjects with high cognitive abilities could utilize the technique (Hill et al., 1997) or it would only help a broader range of subjects if the memory table with the 100 images is present during memorization as reminder (Patton & Lantzy, 1987; Patton, 1986). Some authors argued that it is unlikely people with average or even good memory would ever be able to memorize a table with 100 images to be utilized in the method and only gifted people could achieve the performances of memory athletes (Lieury & Herbst, 2013). The present study proofed these assumptions to be wrong: After a two day course plus about 20 hours of training, a group of subjects with average fluid intelligence compared to students, tested using the BOMAT reasoning task, achieved to more than double their performance in a self-paced digit memorization task with a time limit of five minutes from 22.46 ± 6.83 digits to 50.77 ± 17.15 digits without having any aids like the image table available during the task.

Yet the study is in line with the skeptical results since after instruction, which after all was a two day workshop including practical tasks, the improvement was marginal and not significant (to 26.00 ± 9.90 digits), and only after the following training strongly enhanced performance was achieved. With about 25 hours of training spread over six to eight weeks following the two-day instruction course, the training was still rather limited compared to the long-term training studies of Ericsson and colleagues (Chase & Ericsson, 1982; Ericsson et al., 1980; Richman et al., 1995) and Kliegl´s subjects (Kliegl et al., 1987). Compared to the latter report, in which subjects could not keep using the mnemonic at a pace of 2s/digit, in the digit span task in the present study, 9 out 13 training subjects managed to improve their digit span by at least two digits despite the same pace of 2s/digit. Five of them (more than a third) even managed to reach a digit span of at least 14 or 15 digits at that pace, equaling ceiling performance in this task since no sequences longer than 15 digits had been prepared. This also includes that similar to Kliegl´s subject BB four out of 13 training subjects from the present study could not apply the technique at that pace, despite showing high training gains in the self-paced digit memorization task with a weak correlation ($p_{uncorrected} = .021$, not significant when corrections for multiple testing was applied) between ZVT at pre-training and digit span improvement afterwards, indicating that those with generally slower processing speed could not get up to that pace within the given training time.

High-intelligence or good memory are no preconditions for success in mnemonic training

The present study also rejects the hypothesis that high cognitive abilities and/or an already better memory, maybe even giftedness are necessary to be able to show strong improvements using the phonetic mnemonic (Lieury & Herbst, 2013). Many, but not all, of our subjects were university students. Using the BOMAT, which is normed against students and graduates, it could be seen that the subjects were a totally average sample of students (BOMAT score on IQ scale 101.8 ± 12.7 for training group). Even the worst performer showed an average performance increase over the three tasks of 57% while the average improvement due to retest effects in the control group was just 17.5% ± 2.2%. Improvements in the digit and words self-pace memory tasks were independent of baseline performance in the ZVT and the BOMAT showing that intelligence was not an important factor for benefiting from the mnemonics.

Of note, however, only 13 out of 19 subjects who started the at-home training after taking part in the initial course completed the program. 6 subjects did not comply to or did not finish the training plans and dropped out. One subject did not come to the training course after pre-test. It cannot be ruled out that little or no success in the training was a reason, but group statistics do not support this notion. The drop-outs and those who finished the program did not differ in any important aspect including ZVT score, BOMAT score, and initial memory abilities, expected success and enjoyment of the memory training or motivational type. But there was one highly significant difference in external motivation driving interest in the study: Those who were more motivated by the payment would rather do the training than those who dropped-out. This seems to be plausible since the monetary reward was only paid out to those who finished the whole study and drop-outs received no payment. If one was mostly interested in the memory training course, the reward motivating to join the study was already received in contrast to those eager for the money. Inspection of those who did not finish the program was done:

- Three of the six subjects did less than two hours of training at all and reported various private or study related reasons for not being able to spend the time training including one believable report of a blow of fate regarding private issues.
- One subject reported to actually do the training despite little training automatically logged, and argued the training time reported was not correct due to technical issues.

Log files and implausible reports on technical issues let it seem likely these were false excuses.

- One subject completed 16 hours of training over 8 weeks and managed to gradually improve his performance in the five minute digit memorization task to 86 digits.
- One subject completed little over four hours of training in the program Memocamp and reported some extra hours outside. He asked several questions about the mnemonics after the course and reported having troubles to apply the phonetic mnemonic or seeing any improvement with it.

These reports indicate only one of 19 subjects had actual troubles with the mnemonic techniques that could have influenced him aborting the training, Since he only completed four hours of online training it is impossible to judge whether or not more patience would have eventually allowed him to improve.

So why do some studies see only more intelligent subjects succeed with mnemonics? One possible explanation might be how directly subjects are thought to apply the mnemonic. In the present study the use of the mnemonics for the memory tasks was exactly trained on the same tasks during instruction and during training. An earlier study pointing in this direction showed that all students benefit from mnemonics, but only gifted could transfer them to related memory tasks (Scruggs, Mastropieri, Jorgensen, & Monson, 1986). Also form and amount of instruction might be important since it has been shown that even small variations in instruction can influence performance when mnemonics are taught (Massen et al., 2009). For the method of loci instructions and advices exist for thousands of years (Yates, 1966). Usually the use of well-known real life locations is suggested, the importance of visualization pointed out and learners are instructed to walk along the locations in real life to better store them in memory (Konrad & Dresler, 2007).

During the instructional course in the present study these principles were followed when memorizing and preparing the routes of the method of loci, and shared between the subjects as advice when setting up their own additional routes during training afterwards. Yet some of the studies finding the method to be less effective vary these instructions without giving reasons for this. For example Nyberg et al. (2003) report that some of the older adults (average age just below 70) did not improve from the mnemonic and argue those maybe did not use the method due to general difficulties in generating visual associations and a reduced cognitive resources. In that study the subjects learned about the method while being placed in a PET scanner. 18 locations were presented repeatedly as words shown on screen that could describe objects in a living room. Instructing the subjects to visualize a ball on every location was the training phase. This lacks many of the suggestions on how to learn the method. Additionally the same locations had to be used twice in succession to memorize two different lists of 18 words. Since it has been shown that retroactive interference influence the performance when the method of loci is used (de Beni & Cornoldi, 1988), this also could have a negative effect. All in all, one might argue that based on all these circumstances it is an indicator for the huge robustness of the method that about half of the elderly (and all of the young subjects) did improve their memory performance using the method.

In general for older subjects not complying to the strategy or not using it as instructed have been shown to decrease training-gains (Verhaeghen & Marcoen, 1996). Similarly for children it is known that they do not always benefit from strategies even though they seem to apply it and constant monitoring of proper strategy use is advised. Also more intelligent children are more likely to apply a previously learned strategy themselves without instruction in future tests (Bjorklund, Miller, Coyle, & Slawinski, 1997). For both children and adults, using individual assessment of proper strategy-use showed to be successful (Brehmer et al., 2007). In our study subjects had to explain images used during the instruction course and got suggestion on how to improve them. This could have contributed to the fact that all training subjects did achieve strong training gains but was not contrasted against a group instructed otherwise. In conclusion, when investigating the usefulness of a mnemonic, it is important to make sure good introduction and some form of control of proper strategy use are assured.

Top performances

When looking at the upper end of the performance spectrum, the best performance was 70 words, equaling a pace of 4.3s / word, the best performance ever reported in a scientific study for a serial-word memory task known to the author. In comparison with memory athletes, currently (as of October 1st, 2013) a performance of 70 words memorized in order within five minutes at a memory competition would result in world ranking position 12 (out of over 500 athletes) in this specific discipline[15], providing evidence that even short term training can lead to extraordinary memory performance.

In the digits task (5 min, self-paced), the top three performers recalled 83, 74 and 70 digits equaling a pace of 3.6 s/digit, 4.1s/digit and 4.3s/digit. In the digit-span task at 2s/digit five out of 13 had ceiling performance since no sequences longer than 15 digits were supplied. Wilding and Valentine (1997) had reported on eight participants of the first World Memory Championships in 1990 and three further subjects recognized as individuals with superior memory, who all had to memorize a 6x8 matrix of digits, equaling 48 digits. Only four out of them were faster than 4.3s/digit and only two were faster than 3.6 s/digit. Luria´s famous subject S (Luria, 1968), still one of the most famous cases of superior memory in the psychology literature, took 180 seconds to memorize 48 digits, equaling a speed of 3.75s per digit. So the top three performers of the present study, after only six to eight weeks of training, had achieved a memory skill for digits that would have beaten most of the competitors of the first World Memory Championships and is superior to individuals whose memory was deemed remarkable enough to be studied in single case studies.

3.4.2. Transfer

A series of studies have found promising results indicating that some forms of cognitive training can positively influence other cognitive domains. A potential improvement in fluid intelligence following a working memory training program has been discussed (Jaeggi et al., 2008; Olesen et al., 2004; Redick et al., 2013). These findings had huge impact, but later studies with failed replications reduced the optimism in the field. Various recent studies,

[15] http://www.memocamp.de/deutsch/Bestenliste/Weltrangliste/Disziplin/16/Woerter-5-Min, abgerufen am 01.10.2013

meta-analysis and reviews try to dissolve conflicts and misunderstandings to unveil the true potential and limitations (Brehmer et al., 2012; Buitenweg, Murre, & Ridderinkhof, 2012; A. R. a Conway & Getz, 2010; Gibson, Gondoli, Johnson, Steeger, & Morrissey, 2012; Green et al., 2012; Melby-Lervåg & Hulme, 2013b; Morrison & Chein, 2011; Owen et al., 2010; Redick et al., 2013; Shipstead, Redick, et al., 2012; T. W. Thompson et al., 2013). It has also been suggested that the original assumption that working memory training is a training without strategy use and therefore more a training of underlying the working memory capacity rather than utilizing the existing capacity better (Jaeggi et al., 2008) might be too narrow and strategy use might play a bigger role than assumed and mnemonic training might be more similar to working memory training than postulated (Morrison & Chein, 2011).

Regarding mnemonic training, some studies looking into older subjects found near and far transfer (Ball et al., 2002; Gross & Rebok, 2011; Zelinski, 2009). For school children it has been found that training in various memory strategies also transferred to mental arithmetic and ability to follow instructions(St Clair-Thompson, Stevens, Hunt, & Bolder, 2010). It has not been studied specifically for a particular mnemonic technique how its training influences fluid intelligence or other general cognitive abilities. Studies with limited training in one specific method report a lack of near transfer like in Chase & Ericsson (1982) where the subject heavily improved on the digit span but already in a letter span showed no improvement at all. Therefore a combination of mnemonic strategies trained might be more promising to generate transfer. In the present study subjects trained two complex mnemonic techniques and additionally were instructed to a range of further techniques and the underlying principles of visual imagery and associations. Two transfer measures were studied: Processing speed as measured by ZVT (a trail-making task) and fluid intelligence as measured by BOMAT. Study 1 (see Chapter 2) on memory athletes had revealed a strong superiority of memory athletes on the processing speed measure that even correlates highly with performance in memory sports quickest discipline Speed Cards. Study 2 further supports this finding in that mnemonic training did also transfer to processing speed in the training subjects, who improved significantly more than controls that also showed small retest benefits. On the fluid intelligence measure athletes nominally improved more than controls, but this difference was not significant. Reviews of the working memory training literature suggest rather small effect sizes, thus with 13 subjects per group our study might be underpowered to find transfer on this measure.

As part of the study measures, training subjects showed strong improvement in a working memory measure, the digit span task. Even though Study 1 on the memory athletes indicates these improvements are achieved by direct activation of brain regions involved in long-term memory processing, working memory might act as a moderator and might be implicitly trained as well. In particular, the mnemonics used are based on making visual associations between to-be-remembered items and existing retrieval structures, like the journeys of the method of loci. This indicates a high working memory load since the information has to be hold in focus to be able to make up associations. Therefore the transfer found on processing speed might have been based on working memory training gains rather than the strategies trained. An argument against this notion is found in Study 1 where the memory athletes do not excel in the n-back working memory task but excel in the ZVT. Also the lack of transfer from a digit-span to letter-span found by Chase & Ericsson (1982) supports the assumption that mnemonic strategy training does not alter working memory capacity, but a recent review also question this for working memory training as well (Shipstead, Hicks, et al., 2012).

Regarding our study´s statistical power to judge on transfer effects it has to be mentioned that the control group was passive. Therefore placebo-like effects might have played a role. The same critique was brought up against working memory training, where it is often addressed by having an active control group that does the same training task, but well below ones capacity limit and without adaptation of the tasks´ difficulty. For mnemonic strategies such pseudo training is even harder to realize, since an instruction into the mnemonics cannot be adjusted to be easier. Future studies might compare various forms of training, for example mnemonic training and working memory training, where the training outcomes against active controls are better known. An alternative might also be to develop a pure placebo condition like listening to white noise and telling subjects this would improve memory for words and digits, but it would have to be carefully analyzed if subjects buy into this. There might also be ethical issues to have subjects do a pure fake training for weeks. Since mnemonic training leads to rather obvious improvements in trained memory tasks as seen in the present study and widely accepted in literature, it is questionable if it is possible at all to induce similar improvement expectancy as the mnemonic training subjects might have.

Working Memory Training transfer to fluid intelligence showed to be linear with training time (Jaeggi et al., 2008). The present study only had a pre and a post test of ZVT. Future studies should look into the improvement of ZVT caused by mnemonic training over time.

3.4.3. Subjects valuation of benefits

Besides being known for over two thousand years (Yates, 1966) and shown to work in studies for decades (Bower, 1970b; Roediger, 1980), survey studies show that mnemonic techniques are hardly ever used by students (Soler, María JoseRulz, 1996), general population (Harris, 1980; Intons-Peterson & Fournier, 1986) and even memory researchers (Park, Smith, & Cavanaugh, 1990) who report rather using external aids like writing things down and asking others to remind them than applying mnemonics. Carney & Levin (2008) coined the term “mnemonophobia (i.e., fear of using mnemonics)” and argued “various lingering misconceptions” like supposed benefits only in immediate but not in delayed recall lead to mnemonic techniques being infrequently used despite good arguments in favor of them. A recent review investigating various learning techniques (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013) gives a “low utility” rating to the keyword mnemonic and visual imagery due to assumed high effort and time needed and reduced applicability but only cite studies with limited to no training in these methods. For the keyword mnemonic Dunlosky et al. also argue it is not more beneficial than retrieval practice which needs less effort and less preparation and which is therefore preferred by the authors. They base this statement on Fritz & Morris (2007), who compared the methods and found them to work equally well for learning vocabulary but had little instruction and no training in the keyword mnemonic.

Important, the paper also reports indications on additional benefits when combining both methods suggesting the methods might not be compared as different options but could complement each other. This suggestion is confirmed by two other studies (McKenzie & Sawyer, 1986; Morris et al., 2005) had found that combining both methods leads to best results. Fritz & Morris (2007) also found that despite equal performances, when asked to estimate their performance, subjects perceived the keyword method as less useful but more enjoyable. This finding is in line with another meta-memory study that also showed subjects to misperceive the value of memory strategies (Karpicke, 2009) and studies showing that training a memory strategy is not good enough, but transfer also has to be trained for subjects to utilize it (Hertzog & Dunlosky, 2012).

Further evidence in line with the assumption that improvements are not immediately valued as such by subjects is found in the present study. When asked to judge their memory skill at pretest and posttest, they correctly estimated that their performance in the self-paced tasks will be much better, but when judging their own memory performance in the domains "general", "for digits", "for names and faces" and "for textual information" on a semantic differential scale scored from 1 (bad) to 10 (good), the increase from pretest judgment to posttest judgment was only significant for digits and names. Even on the scale for digit memory ability the score only mildly increased from 5.4 ± 2.0 to 6.7 ± 2.0 with just one subject giving the estimation to be in or above the 90th percentile. Compared to the actual increase in the memory performance achieved, this is a strong underestimation of the truly achieved skill.

3.4.4. fMRI

Few studies exist investigating into the neuronal correlates of mnemonic techniques. The present study intended to look into activation associated with using mnemonic techniques for digits and tasks. Sadly, the number memorization task performed in the scanner proofed to be too difficult for the training subjects. They did significantly improve on number recognition post scanning, but not in free recall or the ordering recall performed during fMRI. Therefore analysis focused on the words condition of the task.

The average effect of condition for the words indicated successful task execution since Broca area associated with language (Geschwind, 1970) and left middle frontal gyrus associated with verbal memory (Petrides, Alivisatos, Meyer, & Evans, 1993; Wagner, 1998) were most significant. However, when comparing the groups post training findings were different than expected. Based on the findings from Maguire et al. (2003) more activation in retrosplenial cortex and hippocampus were expected, but not found. Maybe, despite the already achieved performance gains, strategy use did not establish in the training subjects up to level of memory athletes with many years of training. Kondo et al. (2005) reported additional activations in middle frontal and lingual/cingulate gyri associated with the use of the method of loci, which could not be replicated either. On the contrary, the present study found a significant deactivation in in right inferior parietal lobule / BA 40 (54, -60, 52) in the training subjects compared to controls. This brain region is associated with articular rehearsal in verbal memory tasks (Chen & Desmond, 2005) and therefore it does make sense to be deactivated when switching from a verbal strategy to a visual associative strategy.

The small number of subjects might have limited statistical power to find differences in brain activation. Additional analyses on the data collected could be interesting to investigate for changes in functional connectivity of MTL/Hippocampus with retrosplenial and medial frontal regions but were beyond scope of this thesis.

4. Study 3: Sleep and Cueing effects in the Method of Loci

4.1. Introduction

Memory formation is often described as a three-step process consisting of encoding, consolidation and retrieval. Mnemonics are applied at encoding and made use of at retrieval, but while it can be shown that mnemonics address long-term memory (Chapter 2, Ericsson & Kintsch, 1995) it is an open question as to how they affect memory consolidation, which is understood as the process by which memories are stabilized, thereby enabling long-term retention.

We know that sleep plays an important role in memory consolidation (Stickgold, 2005), but already the question as to which part or properties of sleep are important is much less clear and seems to depend on various factors like which memory system is involved and how recent a memory is (Genzel, Dresler, Wehrle, Grözinger, & Steiger, 2009). During the night we go through several sleep cycles consisting of phases of separable sleep stages that are used to characterize sleep. The sleep stages, as defined by the American Association Sleep Manual (American Academy of Sleep Medicine, 2007), include rapid-eye movement sleep (REM), which is characterized by rapid movements of the eyes, fast EEG and non-REM (NREM). REM comprises about one quarter of the night`s sleep, is more common in the second half of the night and is most associated with memorable dreams. NREM sleep is further split into light sleep (NREM 1) at the beginning of each sleep cycle characterized by shifting of the EEG from alpha waves (8-13 Hz) to theta waves (4-10 Hz), followed by sleep stage 2 (NREM 2) which is characterized by sleep spindles (short burst of oscillatory brain activity in the 10–15 Hz range) and K-complexes (characteristic high voltage complexes in the EEG). The deepest sleep is slow-wave-sleep (SWS or NREM 3) characterized by delta waves (0.5 – 2 Hz), and previously (Rechtschaffen & Kales, 1968) split further into stage 3 and stage 4 based on the amount of delta activity. Today's scoring guidelines consider that

NREM 3 and 4 together represent a single sleep stage (American Academy of Sleep Medicine, 2007).

Early studies provided conflicting evidence as to whether it is SWS (Fowler, Sullivan, & Ekstrand, 1973) or REM (Empson & Clarke, 1970) that is most important for sleep-dependent memory consolidation, where these studies were principally concerned with declarative memory only. Later studies suggested the dual-process hypothesis which contests that there is a dissociation in memory systems where REM sleep is important for consolidation of procedural memory and SWS for consolidation of declarative memory (Plihal & Born, 1997), but reviews show that this dissociation is an over-simplification. Small changes in the details of tasks, such as task difficulty or which subsystem, for example of procedural memory, is involved, already influence in which sleep stage consolidation takes place (Schabus, 2009; Smith, 2001; Vassalli & Dijk, 2009). Via sleep deprivation paradigms it has been shown that procedural memories are still consolidated when only very little REM sleep occurs (Genzel et al., 2009; Rasch, Pommer, Diekelmann, & Born, 2009). It is not necessary to sleep a whole night to receive the beneficial effects of sleep; studies employing short daytime sleeps (naps) frequently find similar benefits on memory performance as are engendered by longer sleeps (Diekelmann, Wilhelm, & Born, 2009; Tucker et al., 2006) and even ultra-short sleeps of just six minutes lead to benefits, albeit less so than longer naps (Lahl, Wispel, Willigens, & Pietrowsky, 2008).

Looking beyond sleep stages, sleep spindles occurring in sleep stage two have also been found to be associated with memory consolidation (Gais, Mölle, Helms, & Born, 2002; Schabus et al., 2004). Interindividual differences in spindle activity exist in humans and are related to general cognitive abilities (Fogel, Nader, Cote, & Smith, 2007; Fogel & Smith, 2011; Schabus et al., 2008). Spindles can be further differentiated by frequency, and thereby categorized into slow (usually 10–13 Hz) and fast spindles (usually 13-15 Hz), with more consistent findings being shown for fast spindles and their role in memory consolidation versus slow spindles. However, knowledge about spindles and memory is not yet consolidated, with different studies looking into different aspects of spindles (for example spindle density, spindle activity and spindle length) or using different criteria to define spindle ranges. So far robust experimental manipulation of spindle activity has not been successful and therefore studies can only speculate as to the causal role of spindles (Rasch & Born, 2013).

Various theories are discussed of how sleep influences memory with views shifting from a previously assumed passive role to more active functioning (Rasch & Born, 2013). One reason for this is that reactivation of memory traces in sleep were found. Initially found in rats, where place cells in the hippocampus fired in the same manner during SWS as they did during a prior learning session in which the animals run along a track (Wilson & McNaughton, 1994); even the temporal order is preserved (Skaggs & McNaughton, 1996). Replay has also been found in REM (Louie & Wilson, 2001). In humans studies using neuroimaging or intracranial recordings in epileptic patients indicate reactivation (Axmacher, Elger, & Fell, 2008; Peigneux et al., 2004), even though direct proof such as the replay of place cells in rodents is hardly possible (Oudiette & Paller, 2012). Further evidence was found when sleepwalker (Oudiette & Constantinescu, 2011) and REM sleep behavior disorder patients (Boeve, 2010) were observed re-enacting recently learned movements during sleep.

A new approach involves triggering or influencing reactivation during sleep by applying external cues. Two recent studies garnered a lot of attention in this field. Firstly a study from 2007 by Rasch, Büchel, Gais, & Born associated learning of a visuospatial task (image-location pairs) with an odor. During the following night, the same odor was presented during SWS. This was compared to a vehicle-only presentation, to odor presentation in REM sleep or in wakefulness, and odor presentation without prior association during learning. Only when the odor was presented during learning and as a cue during subsequent SWS was performance in recall the next morning significantly enhanced (without any odor presentation during recall). Presenting the external cue, in this instance the odor, is thought to have triggered reactivation and thereby strengthened memory consolidation leading to improved recall performance, and fMRI during odor presentation in SWS showed activation of the left hippocampus. The same study did not find a cueing effect of the odor on a procedural memory task, namely finger tapping (Rasch et al., 2007).

Where the Rasch et al. study used one odor as a cue for the whole learning session, another group used specific auditory cues (Rudoy et al., 2009) associated with each item with a similar task. For example, when the spatial location of the image of a cat on a screen had to be learned, the subject heard a corresponding “meow” sound. After learning 50 items subjects took an afternoon nap. Integrated into white noise played during the whole nap, when SWS was reached, 25 sound files associated with learned items were played. After waking subjects were tested on all 50 items and showed significantly improved recall performance for those items that had been cued during the nap. When tested on the various sounds, the subjects reported no awareness that sounds had been played and could not identify the sounds. Thus the Rudoy study suggests that individual items can be strengthened individually by cues presented during sleep. When played during wakefulness, the cues did not have any positive influence. On the contrary, in a later study from the same group, cues were also effective during wake (Oudiette, Antony, Creery, & Paller, 2013). In this later study items were also valued during memorization and the cueing benefit generalized for all low-value items in contrast to the specificity of the study by Rudoy et al.

When replicating the original study design while the subjects slept within the MRI, additional activation was found in the right parahippocampal cortex during cue sounds as compared to control sounds (van Dongen et al., 2012). The authors also found cue-related activity occurred in the bilateral thalamus, cerebellum, and medial temporal lobe correlated with better performance (van Dongen et al., 2012). The cueing effect has also been found for procedural memories (Antony, Gobel, & O’Hare, 2012; Schönauer, Geisler, & Gais, 2013) but when tested on declarative memory only has been shown to exist for spatial tasks (Oudiette & Paller, 2012). Another recent animal study applied auditory cues. Rats learned associations between two auditory stimuli and two sides of a track. Cueing during non-REM sleep led to a bias in reactivation towards the side associated with the sound played and also influenced behavior but did not raise the total number of reactivations, suggesting that cueing biases which information is replayed but does not initiate additional replays (Bendor & Wilson, 2012).

The study reported in this chapter was aimed at finding sleep effects and cueing effects on mnemonic learning using the method of loci. The method of loci enables one to transform a word-list learning task into a visuospatial memory task. Words are visualized as images and placed on locations that had been learned beforehand (see 1.3.7 Method of loci). Therefore applying this method should make the word-list learning task suitable for cueing. Naive subjects had a one-hour introduction session to the method of loci during which two separate sets of 25 locations were learned. They came back three times. On two of the study-days subjects had to learn 50 words using the method of loci and heard fitting sounds for each item. Recall was performed after learning within the MRI. Afterwards subjects either went into the sleep-lab for a nap or stayed awake for the same amount of time. If sleeping, cues were played when a subject got into SWS. Only every second item of one of the two lists was cued, and as such one list remained totally uncued thus allowing for comparisons between both lists as well as within the cued list, where half of the items were cued. A second recall after nap or wake was once again performed in the MRI. The local ethics committee approved the study design. Subjects were paid an honorarium of 100 Euros for participation.

4.2. Methods

4.2.1. Subjects and Design

20 healthy, male subjects were included in the study. Inclusion criteria were right-handedness, no history of psychological or neurological diseases, no substance abuse, no current medication, no experience in mnemonics, German mother tongue, not meeting any exclusion criteria for magnetic resonance scanning, no sleep problems, no shift-work and provision of written informed consent to take part in the study. The mean age of the subjects

was 22.45 ± 2.87 years with an age range between 19 and 30 years. As in study 2, the MWT-B vocabulary test (Lehrl et al., 1995) was used as a screening measure for linguistic ability with a minimum raw score of 18 used as the cut-off mark, since word learning was an essential part of the study. The short depression scale BDI-V (Schmitt et al., 2003) was used as a screening measure for depressive symptoms with a cut-off score of 35 and a standard fMRI screening questionnaire was used to assess fMRI suitability.

The first study day consisted of a structural brain scan, both for screening purposes and to familiarise subjects with the scanner. The screening questionnaires mentioned above, in addition to ZVT as a short measure for processing speed, the VVIQ questionnaire as a measure of visual imagery, a memory self-assessment questionnaire and a questionnaire on motivation based on the Achievement Goal Questionnaire (AVQ) combined with questions on expectancy regarding the memory training were employed. For detailed descriptions of these questionnaires see Chapter 3.2.4.

Afterwards subjects took part in a one-hour introductory session in the method of loci. Two separate lists of 25 locations were learned. The first list consisted of places solely within the institute. The second list consisted of places outside in a nearby park. A third short list of ten locations was taught to provide practise for visualizing and associating words with locations but no images were associated to the study lists during the introductory session. Subjects had to repeat the words learned within the training list to test for their understanding of the method. In comparison to study 2, all subjects thus had the same sets of locations prepared. Subjects were informed that they would be tested on the locations before the following study days and were asked to review the locations, but not to further train in the method. They also had to keep a sleep diary during the study and were asked to maintain regular sleeping habits on the study days.

The actual study consisted of three study days that were undertaken in a random order to prevent sequential effects. On all days subjects came in to the lab in the early afternoon. Afterwards they had to perform a learning task (see below) in condition A and B followed by recall in the MRI. On day A recall was followed by a nap in the sleep laboratory, and on day B it was followed by a wake period where the subjects watched a non-arousing movie. After either their nap or movie, a second recall was performed in the scanner. In condition C, subjects did not undergo a learning session but just took a nap in the sleep lab (see Figure 32). There was a break of at least one week between the two study days including learning to prevent interference since the same sets of locations were used for the method of loci.

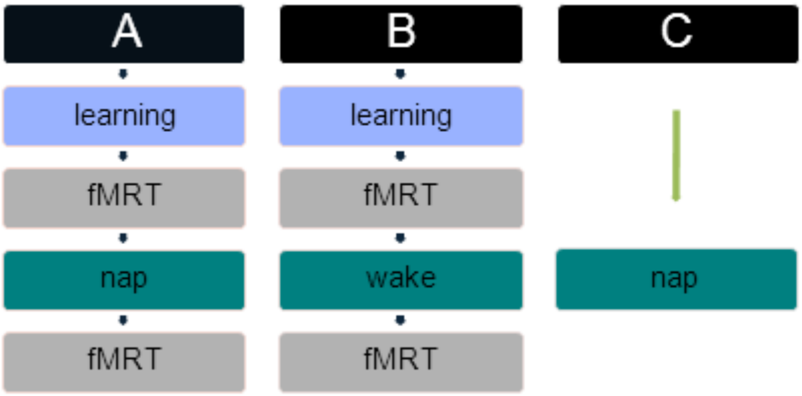


Figure 32: Study design of the three study days that were gone through in random order. There was a minimum break of one week between both days with learning.

4.2.2. Memory task

The main part of the study was the memory task. In their study on acoustic cueing in sleep and learning, Rudoy et al. (2009) used a visuo-spatial learning task in which subjects studied the spatial locations of 50 items on a screen where each item was cued with a characteristic sound (for details, see Rudoy et a. (2009) supplemental material[16]). In the present study, subjects learned 50 words in serial order presented on a screen as text by using the method of loci, i.e. making visual associations with the 50 previously learned locations. Since subjects need to know the locations well enough to go through them mentally without much effort, they were tested on the locations before the task and had to be able to recall them in order without mistakes. If a subject had made a mistake or could not come up with the order,

[16] Retrieved from http://www.sciencemag.org/content/suppl/2009/11/18/326.5956.1079.DC1/Rudoy.SOM.pdf, September 29th, 2013.

he would have been send home, asked to review the locations and come back a different day, but that was not necessary as all subjects could reproduce the locations. Instruction to the whole procedure of the day was given before the start including explaining or reminding of the procedure in the MRI.

A pool of 112 words with matching sound files was built for the task. All words were concrete nouns. For each subject the words were randomly assigned to three sets, 50 words for the study day A, 50 words for study day B and 12 words respectively their sounds were used as control sounds during the nap on study day C.

Learning took place in an office on a 15in TFT screen at a distance of about one meter. The task began with an instruction screen and went on with a button press by the subject when ready. Words were displayed one by one in white font centered on a black background for three seconds each followed by a black screen for one second before the next word. There was a 30 seconds break after the first 25 words to allow subjects to mentally switch to the second list of locations. For each word a matching sound file of about 200ms to 500ms length was played via desktop speakers placed next to the screen with sound onset synchronized with the onset of the word displayed. The volume was normalized for all sound files. After all 50 words were learnt subjects had to recall the words in order using a standard keyboard. Each word was cued by the current location given as current list (“List 1, Institute” or “List 2, Park”; text was displayed in German) and location (1 to 25) as number. Maximum recall time per word was ten seconds. Upper or lower cases were ignored; otherwise the recall had to be correct without spelling mistakes to be counted as correct by the program. If a subject confirmed the current word by pressing the ENTER key or ten seconds passed, the next location was displayed together with a alerting sound to inform of the switch to the next location. A learning criterion of 60% correct was implied for both lists separately, which means at least 15 out of 25 words of each list had to be recalled correctly at the exact serial position. If a subject did not reach the criterion on one or both of the lists, learning of the specific list including sound files played and subsequent recall test repeated until the criterion was reached for both lists. No item-wise feedback was given during recall so the only feedback a subject got was the implicit feedback to have at least 60% of the items of a list correct that was gained by reaching the criterion.

4.2.3. Recall in the MRI

Directly after finishing the learning task subjects were brought to the MRI. Here they again performed a location-cued recall similar to the previous one during the learning task beforehand. Subjects did not type in any answers in the scanner but had a MRI-compatible keyboard with two buttons in their right hand and indicated if they assumed to know an item by pressing the left button or if they assumed not to know the item by pressing the right button. After every third item there was a 10 second break in which a fixation cross was displayed. Subjects were asked not to continue memory recall but take a break with eyes-opened during fixation cross display and only proceed to the next location in their method-of-loci journey when it was cued on screen. The display was projected onto a MRI compatible screen, which the subjects saw via a mirror attached to the head-coil.

After recall in the fMRI, the subjects got a sheet of paper with 50 empty boxes and were asked to write down all the words they could remember in correct order. This recall was taken as measure of performance. Words were counted as correct regardless of obvious spelling mistakes (e.g. ‘Teelöfel’ instead of ‘Teelöffel’) and singular/plural mistakes (‘Katzen’ instead of ‘Katze’) but not if a wrong but similar word was recalled (‘Löffel’ instead of ‘Teelöffel’) or order was wrong (‘Katze, Teelöffel’ instead of ‘Teelöffel, Katze’). Later controls analyses with either stricter or less strict rulings found no effect on the results reported here.

4.2.4. Targeted memory reactivation during nap

After recall on study day A, subjects went to the sleep laboratory, changed to sleeping clothes. EEG electrodes were placed on the scalp (nine electrodes), EOG electrodes next to the eyes (two electrode), EMG electrodes on the chin (three electrodes); ECG was also recorded (two electrodes). The sleep laboratory was located in the basement with no external light or noise. Sleep recordings started approximately 45 minutes after recall in the

MRI. Right from the beginning white noise was played from speakers under the bed at a volume of about 45 decibel. Subjects were told this is to ensure consistent sound level and prevent external disruption.

The sleep EEG was observed online and when deep sleep was observed, the cues were played integrated into white noise. Only the sounds of every second item of only one of the two lists were played with a random selection of the set to be cued, resulting in 12 sound files to be played in one round of cueing (see Figure 33). The cues were played in the order they appeared during the learning task with one sound file played every five seconds. After all items were played, no cues were played for the next minute. If SWS was stable, cues were played up to four times. If cues caused arousals or transition to lighter sleep, the playing of the cues was halted immediately.

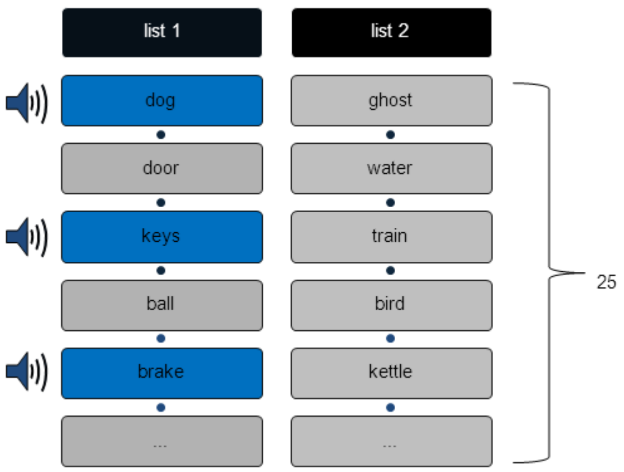


Figure 33: Every second item of one of the two lists, depictured in blue in this figure, was cued by playing its corresponding sounds during SWS with the intention to prompt reactivation of the specific items. Sounds were integrated into background white noise played during the whole nap at about 45dB. A total of 12 items were cued and the cues were played in the same order as the items appeared during learning. After one round of cueing there was a break of 60 seconds. If SWS was stable, cues were played up to four times.

For the analysis of the sleep and cueing effects, only those subjects that reached stable SWS and had all cues played at least two times were included. These were 11 out of the initial 20 subjects. Out of these eleven, eight had the full four runs of cueing, one had three full and one aborted run and two had three full runs. Out of the other nine subjects, eight had no SWS at all, with two subjects not even reaching NREM2.

The nap was ended between 60 and 70 minutes after light-out by opening the door and awakening of the subject. Afterwards electrodes were put off; subjects changed clothes and briefly washed and were then brought back to the scanner for a second recall in the MRI approximately 20 minutes after awakening. They were also asked to fill out a sleep protocol asking if they noted any disturbing noises or other disruptions during sleep. While a few felt distracted by the white noise or thought to have heard an experimenter outside, no one had noticed the cue sounds.

On study day B, subjects did not go to the sleep lab, but were brought to an empty office after the first recall and watched a non-arousing movie on a computer screen for about 120 minutes before returning to the scanner for the second recall.

On study day C, subjects only underwent a nap in the sleep lab without prior learning. To control for effects of the sounds played, 12 items were selected as cue sounds for the control nap that did not appear in the learning task on either day for this subject. Sound files were equally played during SWS only as in the nap following task.

The sleep recordings of the naps were later scored by experienced professional sleep scorers blind to whether and when sound files had been played. They confirmed that for six out of the eleven subjects where cues had been played during the nap following learning, all cues were in SWS. For three subjects most of the cues were played in SWS with some being in NREM 2 and for two subjects all of the cues were played in NREM 2. The further analysis of the sleep data is not part of this thesis and will be discussed elsewhere.

4.2.5. Data analysis

Analysis of behavioral and sleep data was done using SPSS 18. Data is reported as mean ± standard deviation (s.d.) despite where otherwise noted. Significance was assumed for an alpha of 5% and is reported in steps of * = $p < .05$, ** = $p < .005$ and *** = $p < .001$.

The number of repetitions necessary until the learning criterion was reached was taken as first performance measure. The number of items correctly recalled during recall either pre or post nap/wake is the second measure. Additionally a timed recall score was calculated (Lövdén, Brehmer, Li, & Lindenberger, 2012) as the number of correctly recalled items divided by the log of the encoding time in seconds (75s per route) to take into account that subjects might show different scores despite equal number of repetitions.

To compare nap and wake conditions, the retest performance (after nap or wake) was divided by the performance right after finishing learning, resulting in the sleep-dependent consolidation measure presented as percentage. Based on the literature regarding targeted memory activation a beneficial effect of cueing was expected. To analyze this, the performance on cued items divided by the performance of the uncued items was calculated and is presented as percentage. This was repeated 1) for all cued items versus all uncued items regardless of list, 2) all items on the list that contained the cued items versus all items on the other list and 3) only within the list that contained the cued items.

Group comparisons were done using paired t-tests. One-tailed statistics were used where literature suggests the direction of possible effects (sleep effect, cueing effect).

4.2.6. fMRI data acquisition and analysis

fMRI data acquisition and preprocessing equal Study 2. fMRI was carried out at 3 T (Discovery MR750, GE Healthcare, Waukesha, WI,USA) using an 12-channel head coil and covering 42 AC-PC oriented slices (2 mm thickness, 0.5 mm gap; 128 × 128 matrix, interleaved echo planar images, TR 2500 ms, TE 30 ms). fMRI analysis was done with Matlab2008b and SPM8 software and preprocessing was done with the steps slice-time correction, realignment, normalization, reslicing and segmentation. The first four images were discarded after preprocessing to remove non-steady-state effects. See 3.2.6 for more details.

Task effect

Activations and deactivations associated with task-performance were calculated at voxel wise corrected threshold of $p < .05$. Additionally statistical maps were calculated comparing correct and wrong items. The number of wrong items varied between subjects and was very small for some; and in particular due to the applied learning criterion small by definition. Therefore distraction or other forms of unusual behavior that led to misses could influence the results much more than appropriate. Therefore instead of “correct” versus “wrong” items, as actual comparison ”all” versus “wrong” items as suggested by Vanrullen (2011).

Sleep and Cueing effects

Interactions of pre/post sleep and wake as well as pre/post cued and uncued items were calculated to investigate into the neuronal correlates of sleep and cueing effects.

4.3. Results

4.3.1. Behavioral results

20 male subjects were included with a mean age of 22.45 ± 2.87 years and an age range between 19 and 30 years. The cognitive ability was measured with the ZVT and given on an IQ scale the group mean was a score of 108.9 ± 13.6 with a range from 77 to 130.

In the memory tests, the number of repetitions necessary to achieve the learning criterion varied between the subjects. Averaged over both lists and both study days including learning, the mean number of runs necessary were 2.30 ± 0.78 runs. The worst performer was 2.8 SD above the groups range with an average of 4.5 runs and was taken out of the data. He was not included in the sleep and cueing effect data anyway because of lack of SWS in the nap. The performance in the ZVT and the average number of runs correlated significantly, with better performers in the ZVT needing fewer repetitions ($r = -.574$; $p = .01$).

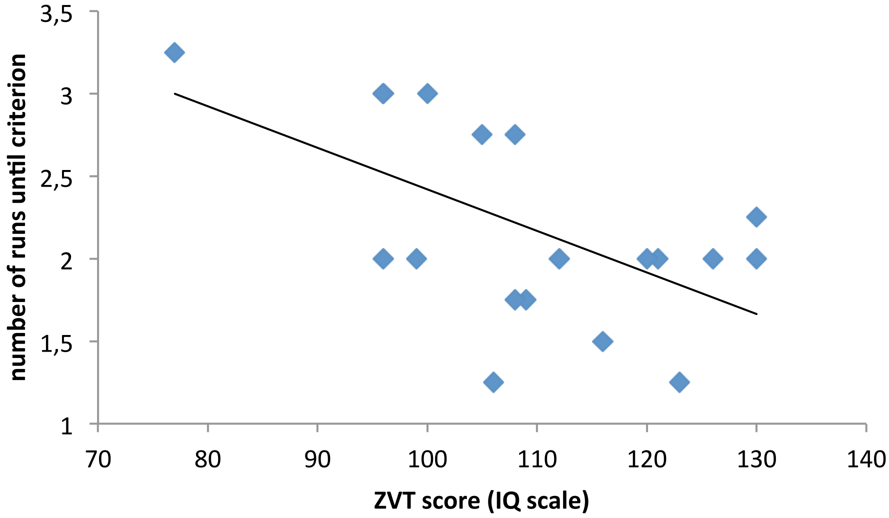


Figure 34: Scatter plot of the average number of runs needed to reach the learning criterion per subject and the performance in the ZVT, a test for processing speed used to assess general cognitive ability.

The average performance on the first free recall test before nap or wake was 42.40 ± 3.91 words correctly recalled at the correct position, equalling a rate of about 85% correct, which is clearly higher than the learning criterion of 60%.

Indoor vs. Outdoor locations

The assumption that outdoor locations work better than indoor locations would result in different recall-performances (Massen et al., 2009) was tested comparing the average number of repetitions needed on both days using paired t-tests but no difference was found ($t(39) = .114$; $p = .910$). For the timed recall (Lövdén et al., 2012) as the number of correctly recalled items divided by the log of the encoding time in seconds no difference was found either ($t(39) = .561$; $p = .578$).

4.3.2. Sleep effect

The following analysis on the sleep and cueing effects is limited to the 11 subjects who reached SWS in the post-learning nap and during which cues were played. Performance post sleep respectively post watching a movie was divided by performance after learning for the performance measure.

Surprisingly, despite no feedback given after the initial recall before sleep, the value is above 100% for several subjects post sleep and for one subject post wake indicating an actual improvement of the recall performance. The mean post-sleep performance was 102.28% ± 3.84%, while the mean post-wake performance was 99.11% ± 1.87%. Based on broad sleep/memory literature a positive effect of sleep was to be expected and a one-tailed paired t-test revealed the difference between both groups to be significant ($t(10)=2.150$; $p < .05$; see Figure 35). The effect size was rather large at Cohen´s d = 1.07. Analysis of the sleep data including correlations of sleep stage duration and number of spindles with memory consolidation is not part of this thesis and will be discussed elsewhere.

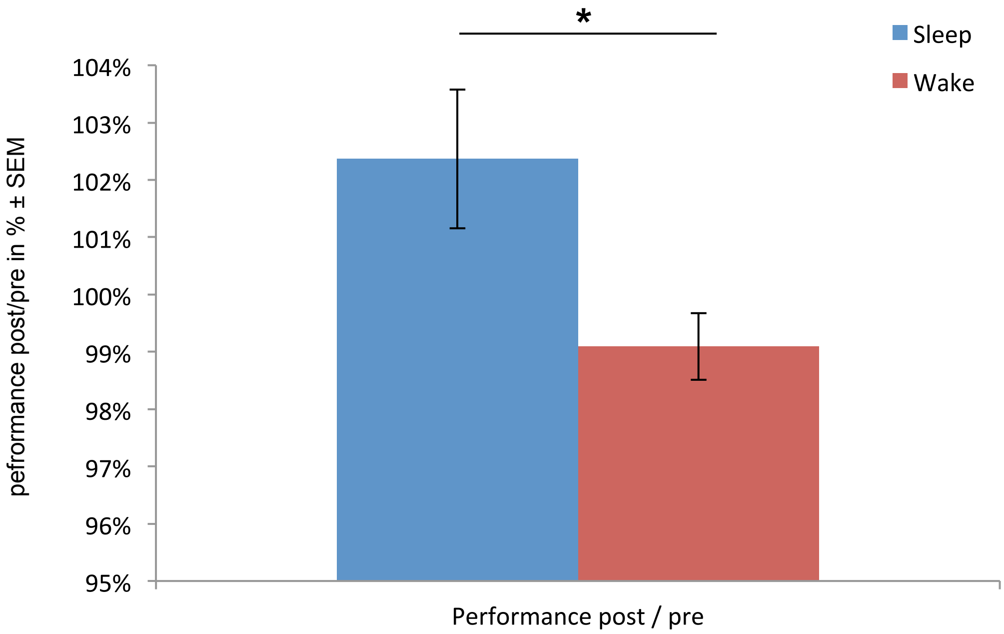


Figure 35: Performance in the free recall of the words after either sleep or wake divided by the performance before sleep or wake showing a significant sleep effect (in percent + SEM). Also when using the method of loci, a nap leads to improved recall compared to a period of wake. * p < .05

4.3.3. Cueing effect

Based on the literature regarding targeted memory activation a beneficial effect of cueing was expected. First, all cued items were compared all uncued items. Since only every second item on one of the two lists was cued, this leads to 12 cued items and 38 uncued items per subject. For the cued items the post/pre performance was 104.24% ± 11.36%, whereas for the uncued items it was 101.99% ± 3.05%, however this difference was not significant in a one-tailed paired t-test (t(10)=.679; p > .05; see Figure 36).

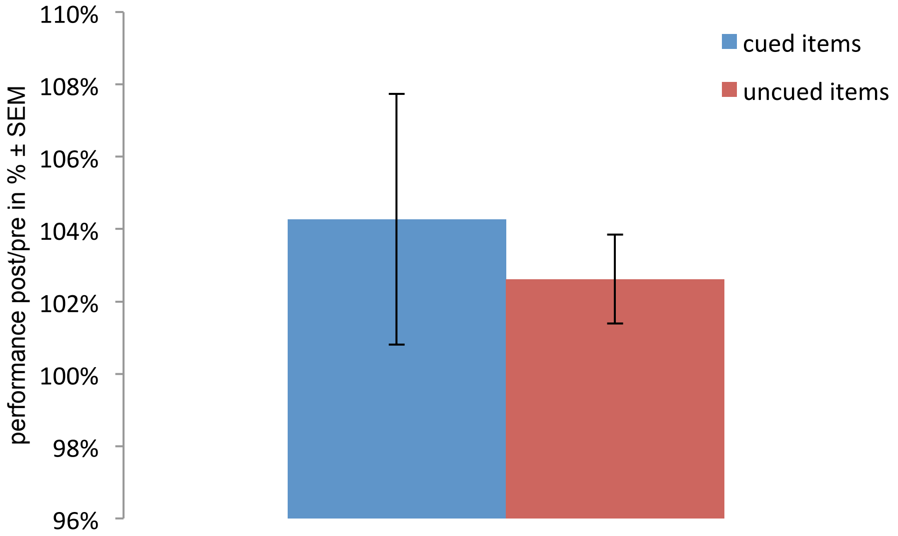


Figure 36: Post / Pre recall performance for all 12 cued items (every second item of one of the lists) versus all 38 uncued items (including 13 items from the list containing the cued items and all 25 items of the second list). The difference was not significant as assessed by a one-tailed paired t-test (p > .05).

A recent study (Oudiette et al., 2013) suggests that the cueing effect might generalize to a whole set of items that are perceived as belonging together. This leads to the alternative hypothesis that the whole list on which every second item was cued could have benefited compared to the second list that remained completely uncued. To test this, post/pre performances were calculated for the cued versus uncued list in every subject. In the cued list the recall performance post/pre was 102.44% ± 8.68%, in the uncued list it was 102.37% ± 4.36%. Obviously this difference is not significant (t(10)=.023; p > .05.

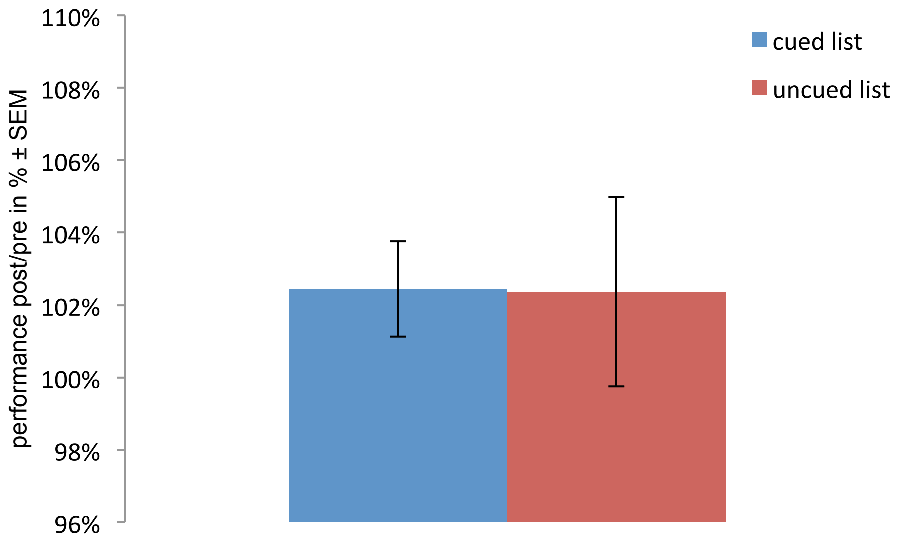


Figure 37: Performance differences post/pre for the two lists. On the cued list 12 out of the 25 items were cued during the nap while the other 13 items were not. The uncued list remained completely without cues. No difference is apparent in the comparison between both lists.

As a final comparison, based on a recent animal study that suggests a biasing effect indicating that cued items would benefit on the costs of uncued related items (Bendor & Wilson, 2012), the cued and uncued items only within the cued list were compared. As mentioned above, the recall performance post/pre for the cued items was 104.24% ± 11.36% whereas for the uncued items from the same list it was 100.81% ± 7.19%. A one-tailed paired t-test revealed this difference to be significant at $p < .05$ ($t(10) = 1.824$; $p < .05$) with a rather small effect size of Cohen´s $d = .35$.

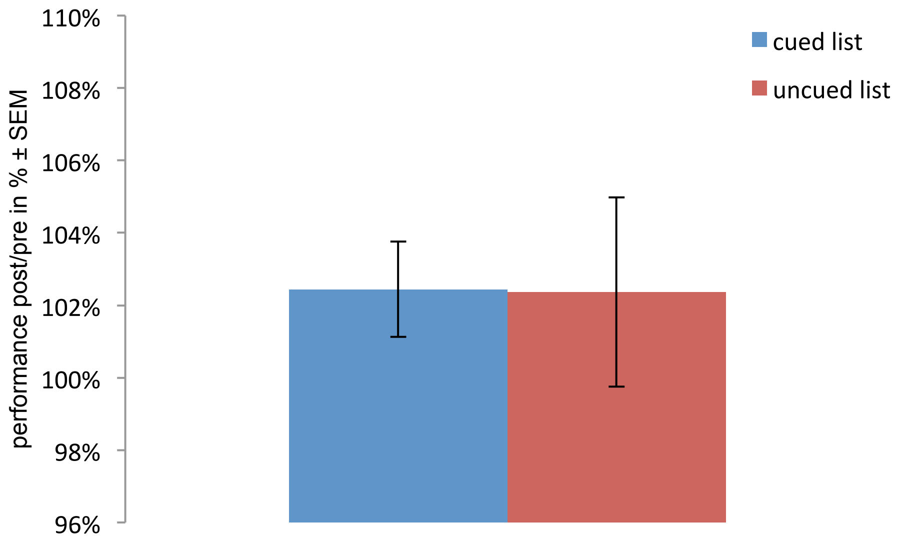


Figure 37: Performance differences post/pre for the two lists. On the cued list 12 out of the 25 items were cued during the nap while the other 13 items were not. The uncued list remained completely without cues. No difference is apparent in the comparison between both lists.

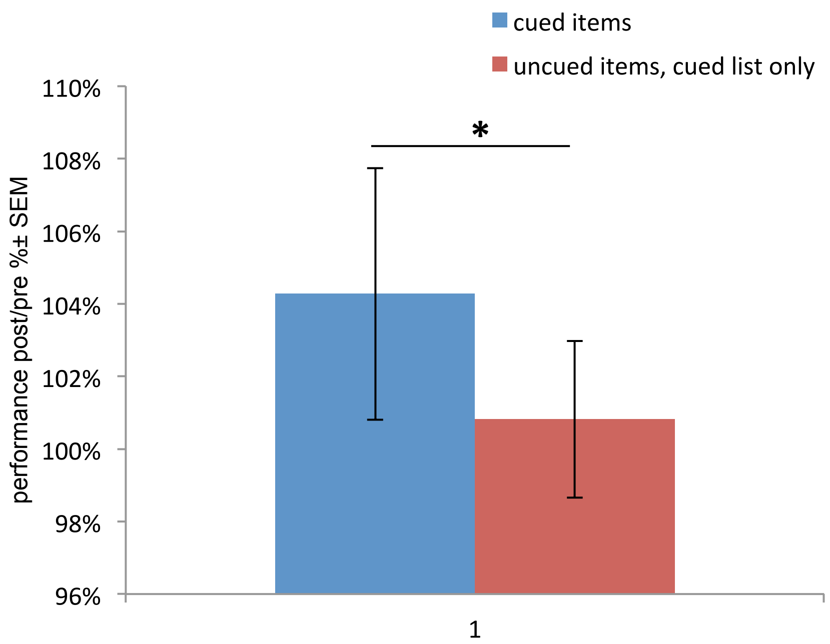


Figure 38: Performance differences post/pre within the list that contains the 12 cued items. The difference between the cued items and uncued items was just significant at p = .049 using a one-tailed paired t-test. * p < .05.

4.3.4. fMRI

The effects of condition for the recall task investigated comined over all time points on both study days (gathered at $p_{FWE} < .05$) revealed a wide range of activations associated with task performance (see Figure 39)) including (MNI coordinates (x,y,z) of peak voxel given) among others: bilateral cingulate gyrus (2, 26, 36), bilateral middle frontal gyrus (52, 34, 30; -32, 50, 8), left inferior parietal lobule (-52, -38, 24), left middle temporal gyrus (-52, -38, 0), right inferior frontal gyrus (40, 20, -12) and a large cluster covering parts of parietal and temporal lobe (-20, -100, 10). Deactivations associated with task performance were found in the Default Mode Network (DMN) including parts of the bilateral parahippocampal gyrus (26, -50, -12; -22 -50 -10), bilateral middle temporal gyrus (40, -78, 14; -40 -70 6) and right precuneus (24, -78, 34).

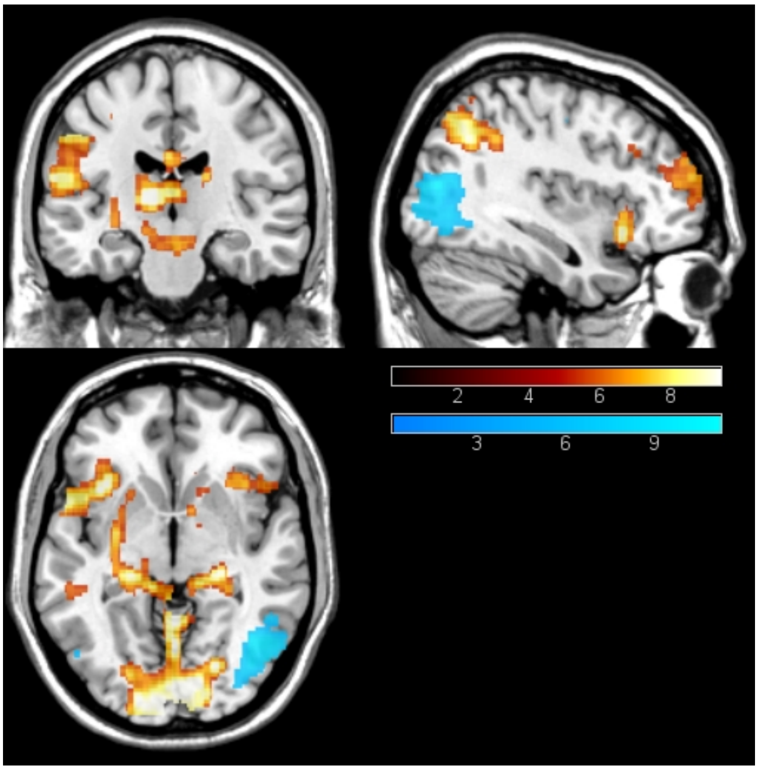


Figure 39: Statistical map for the main effect of condition in the recall task of Study 3 gathered at voxel-wise corrected $p_{FWE} < .05$. Warm colors represent more activation associated with task performance, blue colors deactivation associated with task performance.

When comparing forgotten items to remembered items (Figure 40), it is found that forgotten items correlate with higher activity in bilateral middle occipital / middle temporal gyri (44, -62,

-4; -48, -74, 2), bilateral parahippocampal and fusiform gyri (24, -52, -10; -22, -54, -10) and bilateral precuneus (26, -76, 30; -14, -88, 24).

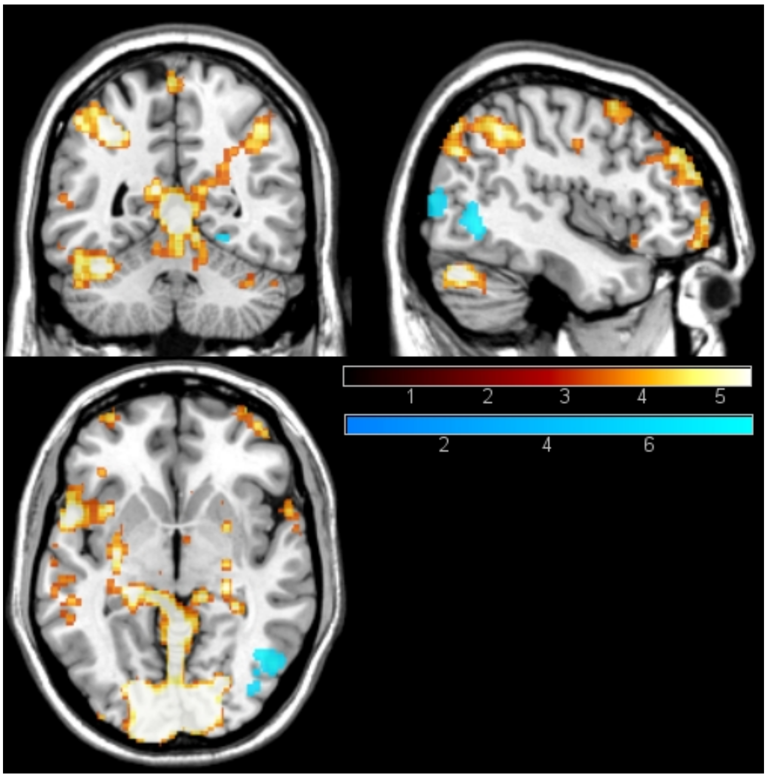
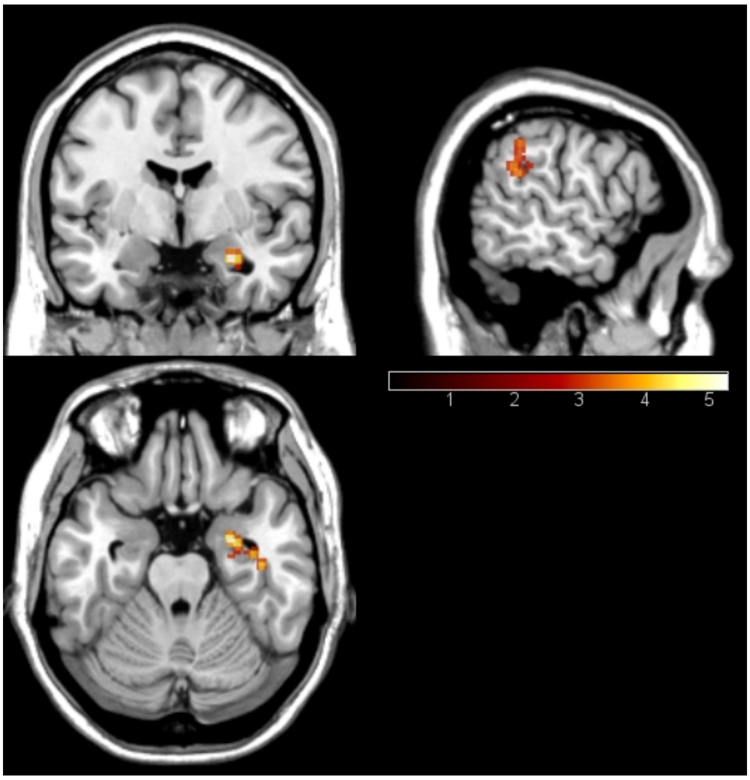


Figure 40: Statistical map for the differences between correctly remembered items and forgotten items, sampled at uncorrected $p < .001$, corrected at cluster level (FWE) with $p < .05$. Warm colors represent increased activation associated with correct answers compared at cool colors associated with wrong answers.

Looking into correlates of the sleep effect did not reveal a significant ($p < .05$, cluster-wise error correction) interaction before and after sleep compared to wake for any regions sampled at uncorrected $p < .005$. Looking into correlates of the cueing effect via the interaction of cued items versus uncued items before and after the nap looking only within the route containing the cued items did reveal two significant clusters: Right parahippocampal gyrus (26,-4,-20) and left inferior parietal lobule (-60, -44, 24), see Figure 41.



.

Figure 41: Interaction pre/post sleep for cued items compares to uncued items sampled at uncorrected p < .001, Right parahippocampal gyrus (26,-4,-20) and left inferior parietal lobule (-60, -44, 24) were found as significant clusters (corrected at cluster-level (p < .05)).

4.4. Discussion

4.4.1. Applying the mnemonic

Since numerous studies demonstrated that the method of loci could be applied on word list learning after some instruction (see Chapter 1.3.7), it was no main goal of the present study to test its efficacy. Nevertheless it is important for the following observations to note that all subjects achieved a 60% learning criterion (i.e. at least 30 words out of 50 recalled at the correct serial position) after only 2.30 ± 0.78 runs with one run equalling each word to be shown for three seconds. In fact, the recall performance was well above 60% at 42.40 ± 3.91 out of 50 words being correctly recalled on average. The instruction in the method took about one hour and included the preparation of two routes of 25 locations each as well as one training round where twenty words were memorized in order using an additional list. No further training followed the instruction session. Since the same locations had to be used at both study days, there was at least one week between both to reduce proactive interference (de Beni & Cornoldi, 1988; Massen & Vaterrodt-Plünnecke, 2006).

The number of runs needed to reach the learning criterion was correlated to the ZVT performance. This is in line with the finding that memory athlete's performance in Speed Cards correlates with ZVT (Study 1) and Digit Span performance correlates with ZVT before and after training (Study 2). It indicates that the cognitive speed of an individual limits his speed in applying mnemonic techniques, but given enough time all subjects can utilize the methods.

The performance of above 40 words in a serial word-list learning task is higher than in other memory studies. A task often used is the paired-associate task, in which words are not learned in serial order but linked to a second word that is later presented as cue. Even with that design often performances only around 20 words are achieved (Plihal & Born, 1997). So using the method of loci enabled subjects to memorize more items than usually tested with easier cued-recall tasks in studies on sleep and memory consolidation. On the other hand it might be the case that mnemonic learning is different to regular learning and therefore applying this technique alters other memory processes as well, which might influence comparability of results obtained from mnemonic learning studies to other forms of learning.

4.4.2. Sleep effect

The present study was the first looking into sleep and memory consolidation for learning with the method of loci. Indeed a sleep effect was found. After a nap subjects recalled more words correctly than after staying awake. It is somewhat surprising that the recall performance after the nap actually is at 102.28% ± 3.84% and by that higher than before the nap. While similar results are sometimes found in paired-associate learning tasks, these usually give feedback after recall (Genzel et al., 2009; Plihal & Born, 1997), which was not done in the present study. It is part of a recent discussion, if regarding declarative memories sleep rather enhances retrieval or protects against forgetting (Rasch & Born, 2013). The present findings suggest that protection against forgetting is not the sole reason, since it could not explain an actual performance improvement. This is in line with a recent study looking into this question using word pair learning: The authors differentiated words being actually better remembered after sleep from words being less forgotten after sleep and also find both, suggesting that probably both aspects do play a role (Fenn & Hambrick, 2013).

4.4.3. Cueing effect

Memory reactivation has been named as a potential mechanism to explain sleep dependent memory consolidation. Targeted memory reactivation (TMR) by applying external cues as a study design caught some attention after first studies showed it to influence memory and enhance recall (Oudiette & Paller, 2012; Rasch et al., 2007; Rudoy et al., 2009).

In the present study, a declarative memory task was used and cues were played mainly during SWS. So far only a few studies looked into declarative memory and TMR and used a visuo-spatial memory task, namely object-location pairs (Oudiette et al., 2013; Rudoy et al., 2009; van Dongen et al., 2012) or sound-word pairing (Fuentemilla et al., 2013). An early study had replayed spoken words associated with a picture series (Tilley, 1979). In additional to declarative memory, reactivation paradigms have also been successfully applied for skill learning (Antony et al., 2012; Schönauer et al., 2013) with melody and movements being associated but cueing skill learning only with a contextual odor cue did not improve memory consolidation (Rasch et al., 2007). Associations of the learned items with the cue has been named as an important factor (Oudiette & Paller, 2012) since purely contextual sounds have failed to improve memory (Donohue & Spencer, 2011). The object-location pair studies all used characteristic sounds as did this study, but so far it has not been investigated if this is actually important or not. Here the object-location pair idea was transferred to a serial word list learning task by applying the method of loci, since this method instructs the subject to visualize the word presented as an object on a location. The items were separated into two lists. A beneficial effect of cueing was found on the cued items compared to the uncued items of the same list. This indicates applying TMR on mnemonically learned items using specific and characteristic acoustic sounds works and influences memory consolidation.

Different sleep stages might be address with TMR. All recent studies addressed mainly SWS as the Rasch et al. study (2007) found cueing to improve declarative memory when applied in SWS but not in REM sleep. Tilley (1979) had played the cues during NREM 2 and REM and found benefits on free recall only after NREM 2 cueing. Regarding cueing in wake conflicting results are found with studies suggesting no effect (Rasch et al., 2007; Rudoy et al., 2009), decrease in performance when aligned with interference (Diekelmann, Büchel, Born, & Rasch, 2011) and increase in performance (Oudiette et al., 2013). In the presented study SWS was addressed and later scoring showed most cues to be played during SWS and a few being in NREM2. The same was true for the Rudoy study. Further studies might more specifically address differences of cueing in NREM2 and SWS to investigate if cueing equally works in both but have to deal with lower arousal thresholds in NREM2 than SWS (Busby, Mercier, & Pivik, 1994).

An important question is what cueing actually does to the sleeping brain and how memory consolidation is influenced. Rudoy et al. (2009) suggested that cued items benefit specifically compared to other items when selectively cued by associated sounds. Oudiette et al. (2013) tested items of low and high value and found the cueing effect to be in particular benefiting the low value items by saving them from forgetting and also found that when some low value items were cued, the positive effect was there for all low value items, even the uncued ones, indicating a generalization effect of the cueing to one set of items. While the Rudoy et al. study does not allow judging if the benefit of the cued items also means harm for the other items because no such comparison was done, the Oudiette et al. study includes a nap without cueing actually finds no harm for the high value items when low value items were cued. In contrast, an animal study suggested that TMR leads to a bias, benefiting the cued information at the expense of other (Bendor & Wilson, 2012), suggesting that reactivation happens spontaneously and TMR does not generate additional reactivations, but influences which items are replayed in those happening anyhow. The present study supports the bias model more than the idea that cueing leads to more reactivation and general performance increase. Comparing the two lists, cued items were better recalled than items of the other list and uncued items of the list containing the cued items were remembered worse, indicating a bias towards the cued items.

4.4.4. Caveats

Since the TMR study design is rather new and still evolving, several details might limit the overall value of the presented study when judging the outcomes but also offer guidance for future studies.

Volume

When using cues that are played during sleep, the volume is of importance. It has been shown that the sleeping brain processes acoustic stimuli, e.g. in a study showing differences in EEG response of the sleeping brain to own name presentation versus other name presentation (Perrin, García-Larrea, Mauguière, & Bastuji, 1999), but auditory cortex responses are reduced in sleep compared to wake (Czisch et al., 2002). To prevent arousals and awakenings by sudden sounds, white noise is played during the whole sleep duration in studies using auditory cues. The cues are then integrated into the white noise. Still, the volume of the sound should not disturb the subjects from sleeping, but if too silent, cues might not be processed by the sleeping brain the way intended. In the present study a few subjects did not reach SWS at all and some reported troubles falling asleep with the white noise background. Also a few arousals at cue onsets were observed and cueing was immediately halted at these instances.

Learning Criterion

Studies investigating sleep effects on memory consolidation usually work with learning criterions. If subjects did not learn enough, there is not much to be supported by sleep and effects might be missed. If on the other hand too much of the information is known, ceiling effects might conceal effects. Also, if overlearning happened, there is a chance that during neither wake nor sleep any forgetting happens (Driskell, Willis, & Copper, 1992) and therefore no sleep effects can be studies on the behavioral outcomes. In studies using the associated word pair tasks a learning criterion of 60% is established (Plihal & Born, 1997) and therefore this has been set as a criterion in the current learning task as well. In the cueing studies using the object-location pair task, a minimum distance is set for each item and items are excluded from further repetitions when recalled within this distance. This was not possible with the loci design, since subjects would have to jump ahead on their journey to specific locations, but that is not possible without high effort since the journeys are prepared for sequential use of the separate locations. Therefore a journey was repeated completely with all 25 locations until the criterion was reached, possibly resulting in overlearning of the items that already had been known after the first runs. The overall very high performance after wake and sleep delays is indicating this has happened. Additionally the way the task was programmed lead to a high number of items being counted as incorrect by the computer program during the learning trials. In particular spelling mistakes and missing a location led to inappropriately poor scores and therefore additional repetitions where the subject actually already met the criterion. In the test after learning subjects on average recalled 85% of the items instead of about 60% intended. A further issue was that our testing procedure included retrieving all item learned twice, once in the scanner and a second time for written free recall. Since retrieval practice is known to increase retention (Roediger & Karpicke, 2006), this probably also further consolidated memories and reduced the variability of the data. Future studies intending to utilize the task described should therefore change the procedure in these details. A possible alternative would be to cue the locations during recall in both learning and testing by showing a picture of the location instead of just giving it´s number. Then retrieval can be done in random order besides using the method of loci and items known before can be skipped. If the performance during the last learning round before reaching the criterion is taken as pretest measure, two rounds of retrieval could be saved in comparison to the design used.

Group Size and Power

20 subjects were included in the present study but only eleven reached the necessary amount of SWS to play the auditory cues often enough. The effect size for the sleep effect is large at Cohen´s d = 1.07. A post-hoc power test using G*Power 3[17] reveals a power of .87 to find the sleep effect of that effect size. The effect size for the cueing effect on the cued lists is small at Cohen´s d = .35. The group size was comparable to other studies applying TMR with acoustic cues (Rudoy et al.: 12 out of 17 cued; Oudiette et al.: 15 per group, cued in NREM 2 when no SWS observed; Fuentemilla et al.: Groups of 4, 7 and 9). Yet for future studies replications with larger groups are suggested, since the cueing effects are somewhat smaller and less robust than the sleep effects.

4.4.5. fMRI

The task activation confirms actual task engagements by the subjects as brain areas commonly associated with memory tasks were activated. Since all subjects have used the method of loci, no brain activation can be directly associated with strategy use.

However, when looking into items subject forgot, higher activation was found in parahippocampal areas as well as bilateral middle occipital / middle temporal gyri, which are associated with higher visual processes and visual memory including object recognition. Both areas fit to the assumption a subject used the method of loci but did not find the object stored there. If a subject thinks of the location, where the association was stored, spatial and navigational brain activity should be stronger than if he immediately found the image he was looking for, fitting to higher parahippocampal and fusiform activity. He will also engage in more intense visualizing of the location and trying to identify the objects he stored there, hence activating brain areas associated with visualizing and object recognition.

Looking at the imaging findings for the sleep and cueing effects provide mixed in sight. It was to expect to find different activation after sleep than after week (Rasch & Born, 2013), but we

[17] http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/

did not. On the contrary comparing cued and uncued items within the cued route revealed a significant difference in right parahippocampal gyrus (26,-4,-20) activation associated with cueing during sleep. The parahippocampal gyrus plays an important role in moderating newly learned information and its connectivity to various brain areas is influenced by cueing (van Dongen et al., 2012). A change in its activation post cueing for the items on the cued list supports the behavioral finding of a biasing effect rather than a cueing effect that generalizes towards a whole set of items.

5. Outlook

This thesis presented three studies that looked into superior memory from different angles. Study 1 investigated memory athletes and found that superior memory performance correlated with processing speed and intelligence. Since Study 2 found transfer on processing speed but not on intelligence after only six weeks of training, it would be interesting for future studies to follow a longer memory training interventions as well as training programs of different lengths to further investigate gains achieved on the transfer task.

Since Study 2 only had a passive control group and only one task per constructed investigated, in the light of the current training literature it would be interesting to carefully design a study comparing mnemonic training and working memory training. While the studies presented in this thesis clearly indicate that mnemonic training does not improve working memory, it did lead to the ability to master some tasks classified as working memory tasks by applying mnemonic techniques and arguably utilize long-term working memory for these. On the other hand, memory athletes could not apply their methods in the n-back task they were not accustomed to. Some suggested they might be able to adapt the strategy to also master this task, which is also used as a training task in working memory studies and is argued as unsuitable for strategy use. Therefore it would be interesting to have some athletes interested in trying to train for this task and see how quickly they achieve superior performance in it. In regards to memory performance achievable Study 2 proved that performance levels so far only achieved in training studies on individuals can be achieved in a group setting. Feedback from individual participants and questionnaire answers proved most of them enjoyed the training and were interested in continuing. Thereby besides the possibility to invite memory athletes to participate in studies, it is possible to produce a group of superior memorizers rendering the assumption of not having the possibility to gather groups of superior memorizers for studies (Ericsson, 2003) wrong.

Study 3 is an example of how such groups could be investigated into to answer open questions in memory research. Applying the method of loci on a word-list learning task allowed making word-list learning applicable for a study on targeted memory reactivation (TMR). Besides the limitations of the previous studies as mentioned in the discussion of Study 3, it generated promising results and motivates further applying this mnemonic paradigm on sleep studies. In particular the finding that TMR also biases content rather than promoting additional reactivation is important for the discussion of the possibilities and promises of TMR.

6. *Appendix*

Memory self-appreciation questionnaire, study 2

Selbsteinschätzung

Name: ______________________

Bitte bewerten Sie Ihre Gedächtnisleistung. Bitte kennzeichnen Sie Ihre Einschätzung durch einen senkrechten Strich auf den Linien. Je weiter Sie diesen Strich nach rechts setzen, umso mehr trifft die rechte Eigenschaft zu; je weiter Sie den Strich nach links setzen, umso mehr trifft die linke Eigenschaft zu.

A1) Ihre Gedächtnisleistung (allgemein)

Schlecht |————————————| Gut

B1) Ihre Leistung beim Einprägen von Zahlen

Schlecht |————————————| Gut

C1) Ihre Leistung beim Einprägen von Namen & Gesichtern

Schlecht |————————————| Gut

D1) Ihre Leistung beim Lernen schriftlich präsentierter Informationen

Schlecht |————————————| Gut

A2) Ihre Gedächtnisleistung (allgemein) im Vergleich zu ihrem persönlichen Bekanntenkreis

Unterdurchschnittlich |————————————| Überdurchschnittlich

B2) Ihre Leistung beim Einprägen von Zahlen im Vergleich zu ihrem persönlichen Bekanntenkreis

Unterdurchschnittlich |————————————| Überdurchschnittlich

C2) Ihre Leistung beim Einprägen von Namen & Gesichtern im Vergleich zu ihrem persönlichen Bekanntenkreis

Unterdurchschnittlich |————————————| Überdurchschnittlich

D2) Ihre Leistung beim Lernen schriftlich präsentierter Informationen im Vergleich zu ihrem persönlichen Bekanntenkreis

Unterdurchschnittlich |————————————| Überdurchschnittlich

Selbsteinschätzung, Seite 2

Name: ________________________

Bitte schätzen Sie:

Wie viele Ziffern können Sie sich in 5 Minuten
<u>in der richtigen Reihenfolge</u> einprägen? __________

Wie viele Worte können Sie sich in 5 Minuten
<u>in der richtigen Reihenfolge</u> einprägen? __________

Bitte beurteilen Sie:

Wie viele Ziffern in 5 Minuten <u>in der richtigen Reihenfolge einprägt</u>, sind Ihrer Meinung nach eine außergewöhnlich gute Gedächtnisleistung? __________

Wie viele Worte in 5 Minuten <u>in der richtigen Reihenfolge einprägt</u>, sind Ihrer Meinung nach eine außergewöhnlich gute Gedächtnisleistung? __________

Motivational Questionnaire based on AVQ plus questions on expectancy

Name ________________________

Sie finden im Folgenden 17 Aussagen. Bitte geben Sie auf einer Skala von 1 – trifft überhaupt nicht auf mich zu bis 7 – trifft vollständig auf mich zu an, wie Sie die Aussagen auf Sie persönlich zutreffen. Bitte beziehen Sie Ihre Antworten auf den gesamten Kurs, also das Wochenendseminar und das anschließende Training zu Hause.

01	Ich erwarte, dass ich mich in dem Kurs verbessern werde	1	2	3	4	5	6	7
02	Es ist mir wichtig besser abzuschneiden als andere	1	2	3	4	5	6	7
03	Ich befürchte, dass ich das Gedächtnistraining nicht so gut bewältige wie ich vielleicht könnte	1	2	3	4	5	6	7
04	Ich will mich in den Aufgaben so weit wie möglich verbessern	1	2	3	4	5	6	7
05	Ich möchte nur vermeiden, in den Aufgaben schlecht abzuschneiden	1	2	3	4	5	6	7
06	Es ist mir wichtig in diesem Kurs im Vergleich zu den anderen gut abzuschneiden	1	2	3	4	5	6	7
07	Ich mache mir Sorgen, dass ich mich nicht so sehr verbessere wie ich gerne würde	1	2	3	4	5	6	7
08	Es ist mir wichtig in diesem Kurs so viel wie möglich zu lernen	1	2	3	4	5	6	7
09	Mein Ziel in diesem Kurs ist es nicht schlecht abzuschneiden	1	2	3	4	5	6	7
10	Ich erwarte, dass mir der Kurs Spaß machen wird	1	2	3	4	5	6	7
11	Mein Ziel in diesem Kurs ist es bessere Ergebnisse als die meisten anderen zu erzielen	1	2	3	4	5	6	7
12	Ich bin besorgt darüber, dass ich nicht alles lerne, was man in diesem Kurs lernen kann	1	2	3	4	5	6	7
13	Ich habe das Verlangen die Gedächtnistechniken vollständig zu beherrschen	1	2	3	4	5	6	7
14	Die Angst schlechte Leistungen zu bringen ist oft das, was mich motiviert	1	2	3	4	5	6	7
15	Ich erwarte, dass ich einer der besten in diesem Kurs sein werde	1	2	3	4	5	6	7
16	Ich habe mich vor allem wegen des Gedächtnistrainings für die Studie gemeldet	1	2	3	4	5	6	7
17	Ich habe mich vor allem wegen des Geldes für die Studie gemeldet	1	2	3	4	5	6	7

VVIQ, self-translated version

Zur Beurteilung der Einträge 1 bis 4 denken Sie bitte an einen Bekannten oder Verwandten, den Sie regelmäßig sehen, der aber derzeit nicht anwesend ist. Betrachten Sie das Bild, das vor Ihrem inneren Auge entsteht, sorgfältig.

0	*Vorab: Welcher Wert steht für vollkommen klare Vorstellung?*		
1	Der exakte Umriss von Gesicht, Kopf, Schultern und Körper.		
2	Typische Posen des Kopfes, Körperhaltungen etc.		
3	Die präzise Körperhaltung, Schrittlänge etc. beim Gehen		
4	Die verschiedenen Farben in häufig getragenen Kleidungsstücken		

Stellen Sie sich die aufgehende Sonne vor. Betrachten Sie das Bild, das vor Ihrem inneren Auge entsteht, sorgfältig.

5	Die Sonne geht über den Horizont auf in einen diesigen Himmel.		
6	Der Himmel klart auf und die Farbe blau umschließt die Sonne.		
7	Wolken. Ein Sturm kommt, es blitzt.		
8	Ein Regenbogen erscheint.		

Denken Sie an die Vorderseite eines Geschäfts, das Sie oft aufsuchen. Betrachten Sie das Bild, das vor Ihrem inneren Auge entsteht, sorgfältig.

9	Die allgemeine Erscheinung des Geschäfts von der anderen Straßenseite		
10	Ein Schaufenster incl. Farben, Form und Details ausgestellter Produkte.		
11	Sie sind jetzt vor dem Eingang. Farbe, Form und Details der Tür		
12	Sie betreten das Geschäft und gehen an die Kasse. Die Kassiererin bedient Sie, Geld wechselt die Hände.		

Zum Schluss denken Sie an eine Landschaftsszene mit Bäumen, Bergen und einem See. Betrachten Sie das Bild, das vor Ihrem inneren Auge entsteht, sorgfältig.

13	Die Umrisse der Landschaft		
14	Farben und Form der Bäume		
15	Farben und Form des Sees		
16	Ein starker Wind bläst über die Bäume, erzeugt Wellen im See		

Questionnaire on the memory training after at-home training

Studie Gedächtnistraining
Angaben zum Training

Name: ____________________

Datum: ____________________

Wie viele Stunden haben Sie Ihrer Einschätzung nach mit der Internet-Plattform Memocamp trainiert? [ ]

Wie viele Stunden haben Sie Ihrer Einschätzung nach „Offline" trainiert? [ ]

Bitte beschreiben Sie, in welcher Art und Weise und in welchem Umfang Sie sich nach dem Seminar mit Gedächtnistechniken befasst oder diese außerhalb des Memocamps trainiert haben:

____

____

____

____

Bitte bewerten Sie Ihre Motivation zum Training:

1) Das Gedächtnistraining hat mir Spaß gemacht

Trifft gar nicht zu |————————————————| Trifft voll zu

2) Ich musste mich zum Training zwingen

Trifft gar nicht zu |————————————————| Trifft voll zu

3) Das Training war für mich nützlich

Trifft gar nicht zu |————————————————| Trifft voll zu

4) Ich werde das Gedächtnistraining vermutlich fortsetzen

Trifft gar nicht zu |————————————————| Trifft voll zu

List of words used in Study 3 combined with matching sounds of .2 - .5s duration

Lernbegriffe			
Stammestanz	Bellen	Toilette	Pistole
Kassette	Ufo	Harfe	Wolf
Reisverschluss	Schritte	Scherbe	Rakete
Militär	Vogel	Kartoffelchips	Fotokamera
Bremse	Hahn	Hund	Pferde
Ziege	Regen	Hexe	Rassel
Schwein	Dampfer	Becken	Bauernhof
Auto	Wasserglas	Geist	Säge
Turbine	Hamster	Uhr	Apfel
Bongos	Würfel	Schwert	Schere
Medikament	Indianer	Tram	Schwimmbad
E-Mail	Applaus	Ente	Schlaf
Lachen	Herzschlag	Luftballon	Grille
Cartoon	Weinen	Kuckucksuhr	Affe
Schneebesen	Schluckauf	Reiter	Münze
Zug	Helikopter	Katze	Kaffeemaschine
Lautsprecher	Windspiel	Mücke	Motorrad
Scharnier	Kuhwiese	Feuerwerk	Schafsherde
Eiswürfel	Motorboot	Taubenschlag	Korkenzieher
Tropfen	Buchseite	Gitarre	Klavier
Luftpumpe	Telefon	Anruf	Glocke
Trommel	Peitsche	Paketband	Weinglas
Bohrer	Baby	Coladose	Taschentuch
Geschirr	Tür	Adler	Bierflasche
Klopfen	Teelöffel	Tastatur	Blasen

7. References

American Academy of Sleep Medicine. (2007). *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine.

Anschutz, L., Camp, C. J., Markley, R. P., & Kramer, J. J. (1987). Remembering mnemonics: a three-year follow-up on the effects of mnemonics training in elderly adults. *Experimental aging research*, *13*, 141–3.

Antony, J. W., Gobel, E., & O'Hare, J. (2012). Cued memory reactivation during sleep influences skill learning. *Nature Neuroscience*, *15*, 1114–1116.

Atkinson, R. C. (1974). Mnemotechnics in second-language learning. *American Psychologist*, *30*, 821–828.

Axmacher, N., Elger, C. E., & Fell, J. (2008). Ripples in the medial temporal lobe are relevant for human memory consolidation. *Brain*, *131*, 1806–17.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews. Neuroscience*, *4*, 829–39.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., … Willis, S. L. (2002). Effects of Cognitive Training Interventions With Older Adults, *288*, 2271–2281.

Baltes, P. B., & Kliegl, R. (1992). Further testing of limits of cognitive plasticity: Negative age differences in a mnemonic skill are robust. *Developmental Psychology*, *28*, 121–125.

Baron-Cohen, S., Bor, D., Billington, J., Asher, J., Wheelwright, S., & Ashwin, C. (2007). Savant memory in a man with colour form-number synaesthesia and asperger. *Journal of consciousness studies*, *14*, 9–10.

Belleville, S., Gilbert, B., Fontaine, F., Gagnon, L., Ménard, E., & Gauthier, S. (2006). Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. *Dementia and geriatric cognitive disorders*, *22*, 486–99.

Bellezza, F. S. (1981). Mnemonic Devices: Classification, Characteristics, and Criteria. *Review of Educational Research*, *51*, 247–275.

Bellezza, F. S., Six, L. S., & Phillips, D. S. (1992). A mnemonic for remembering long strings of digits. *Bulletin of the Psychonomic Society*, *30*, 271–274.

Bendor, D., & Wilson, M. a. (2012). Biasing the content of hippocampal replay during sleep. *Nature neuroscience*, *15*, 1439–1444.

Bird, S. A., & Jacobs, G. M. (1999). An examination of the keyword method: How effective is it for native speakers of Chinese learning English. *Asian Journal of English Language Teaching*, *9*, 75–97.

Bjorklund, D. F., Miller, P. H., Coyle, T. R., & Slawinski, J. L. (1997). Instructing Children to Use Memory Strategies: Evidence of Utilization Deficiencies in Memory Training Studies. *Developmental Review*, *17*, 411–441.

Boeve, B. F. (2010). REM sleep behavior disorder: Updated review of the core features, the REM sleep behavior disorder-neurodegenerative disease association, evolving concepts, controversies, and future directions. *Annals of the New York Academy of Sciences*, *1184*, 15–54.

Bor, D., Billington, J., & Baron-Cohen, S. (2007). Savant memory for digits in a case of synaesthesia and Asperger syndrome is related to hyperactivity in the lateral prefrontal cortex. *Neurocase*, *13*, 311–9.

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*, 361–7.

Bor, D., & Owen, A. M. (2007). A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory. *Cerebral cortex (New York, N.Y. : 1991)*, *17*, 778–86.

Bottiroli, S., Cavallini, E., & Vecchi, T. (2008). Long-term effects of memory training in the elderly: a longitudinal study. *Archives of gerontology and geriatrics*, *47*, 277–89.

Bower, G. H. (1970a). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, *9*, 529–533.

Bower, G. H. (1970b). Analysis of a Mnemonic Device: Modern psychology uncovers the powerful components of an ancient system for improving memory. *American Scientist*, *58*, 496–510 CR – Copyright © 1970 Sigma Xi, The S.

Bower, G. H., & Clark, M. C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science*, *14*, 181–182.

Brehmer, Y., Li, S., Müller, V., Oertzen, T. Von, & Lindenberger, U. (2007). Memory plasticity across the life span: uncovering children's latent potential. *Developmental Psychology*, *43*, 465– 478.

Brehmer, Y., Li, S., Straube, B., Stoll, G., Oertzen, T. Von, Müller, V., & Lindenberger, U. (2008). Comparing Memory Skill Maintenance Across the Life Span : Preservation in Adults , Increase in Children. *Psychology and Aging*, *23*, 227–238.

Brehmer, Y., Westerberg, H., & Bäckman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Frontiers in human neuroscience*, *6*. doi:10.3389/fnhum.2012.00063

Bruce, D., & Clemons, D. M. (1982). A test of the effectiveness of the phonetic (number-consonant) mnemonic system. *Human Learning*, *1*, 83–93.

Bugelski, B. R., Kidd, E., & Segmen, J. (1968). Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology*, *76*.

Buitenweg, J. I. V, Murre, J. M. J., & Ridderinkhof, K. R. (2012). Brain training in progress: a review of trainability in healthy seniors. *Frontiers in human neuroscience*, *6*, 183.

Busby, K. A., Mercier, L., & Pivik, R. T. (1994). Ontogenetic variations in auditory arousal threshold during sleep. *Psychophysiology*, *31*, 182–188.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of cognitive neuroscience*, *12*, 1–47.

Campos, A., Amor, A., & González, M. A. (2004). The importance of the keyword-generation method in keyword mnemonics. *Experimental psychology*, *51*, 125–31.

Canellopoulou, M., & Richardson, J. T. (1998). The role of executive function in imagery mnemonics: evidence from multiple sclerosis. *Neuropsychologia*, *36*, 1181–8.

Carney, R. N., & Levin, J. R. (1994). Combining Mnemonic Strategies to Remember Who Painted What When. *Contemporary Educational Psychology*, *19*, 323–339.

Carney, R. N., & Levin, J. R. (2008). Conquering mnemonophobia, with help from three practical measures of memory and application. *Teaching of Psychology*, *35*, 176–183.

Carney, R. N., & Levin, J. R. (2012). Facing Facts: Can the Face-Name Mnemonic Strategy Accommodate Additional Factual Information? *The Journal of Experimental Education*, *80*, 386–404.

Carney, R. N., Levin, J. R., & Stackhouse, T. (1997). The face–name mnemonic strategy from a different perspective. *Contemporary educational …*, *22*, 399–412.

Carretti, B., Borella, E., & De Beni, R. (2007). Does Strategic Memory Training Improve the Working Memory Performance of Younger and Older Adults? *Experimental Psychology (formerly "Zeitschrift für Experimentelle Psychologie")*, *54*, 311–320.

Chase, W. G., & Ericsson, K. A. (1981). Skilled memory. *Cognitive skills and their acquisition*, 141–189.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G H Bower (Ed.), *The psychology of learning and motivation* (16th ed., Vol. 16, pp. 1–58). Academic Press, New York.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: training and transfer effects with a complex working memory span task. *Psychonomic bulletin & review*, *17*, 193–9.

Chen, S. H. A., & Desmond, J. E. (2005). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. *NeuroImage*, *24*, 332–8.

Chooi, W.-T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*, 531–542.

Cohen, G. (1990). Why is it difficult to put names to faces? *British Journal of Psychology*, *81*, 287–297.

Cohen, G., & Faulkner, D. (1986). Memory for proper names: Age differences in retrieval. *British Journal of Developmental Psychology*, *4*, 187–197.

Conway, A. R. a, & Getz, S. J. (2010). Cognitive ability: does working memory training enhance intelligence? *Current biology : CB*, *20*, R362–4.

Conway, A. R. A., Kane, M. M. J., & Al, C. E. T. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*, 769–786.

Czisch, M., Wetter, T. C., Kaufmann, C., Pollmächer, T., Holsboer, F., & Auer, D. P. (2002). Altered processing of acoustic stimuli during sleep: reduced auditory activation and visual deactivation detected by a combined fMRI/EEG study. *NeuroImage*, *16*, 251–8.

De Beni, R., & Cornoldi, C. (1988). Does the repeated use of loci create interference? *Perceptual and Motor Skills*, *67*, 415–418.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature reviews. Neuroscience*, *11*, 114–26.

Diekelmann, S., Büchel, C., Born, J., & Rasch, B. (2011). Labile or stable: opposing consequences for memory when reactivated during waking and sleep. *Nature neuroscience*, *14*, 381–6.

Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep medicine reviews*, *13*, 309–21.

Donohue, K. C., & Spencer, R. (2011). Continuous re-exposure to environmental sound cues during sleep does not improve memory for semantically unrelated word pairs. *Journal of Cognitive Education and Psychology*, *10*, 167–177.

Drevenstedt, J., & Bellezza, F. S. (1993). Memory for self-generated narration in the elderly. *Psychology and aging*, *8*, 187–96.

Driskell, J. E., Willis, R. P., & Copper, C. (1992). Effect of overlearning on retention. *Journal of Applied Psychology*, *77*, 615–622.

Dunlosky, J., Rawson, K. a., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*, 4–58.

Elliot, A. J., & McGregor, H. a. (2001). A 2 X 2 achievement goal framework. *Journal of personality and social psychology*, *80*, 501–19.

Empson, J. A., & Clarke, P. R. F. (1970). Rapid Eye Movements and Remembering. *Nature*, *227*, 287–288.

Engvig, A., Fjell, A. M., Westlye, L. T., Moberget, T., Sundseth, Ø., Larsen, V. A., & Walhovd, K. B. (2010). Effects of memory training on cortical thickness in the elderly. *NeuroImage*, *52*, 1667–76.

Ericsson, K. A. (2003). Exceptional memorizers: made, not born. *Trends in cognitive sciences*, *7*, 233–235.

Ericsson, K. A., Chase, W. G., & Faloon, S. (1980). Acquisition of a memory skill. *Science (New York, N.Y.)*, *208*, 1181–2.

Ericsson, K. A., Delaney, P. F., Weaver, G., & Mahadevan, R. (2004). *Uncovering the structure of a memorist's superior "basic" memory capacity. Cognitive Psychology* (Vol. 49, pp. 191–237).

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review*, *102*, 211–45.

Ericsson, K. A., & Staszewski, J. J. (1989). Skilled Memory and Expertise: Mechanisms of Exceptional Performance. *Complex Information Processing The Impact of Herbert A Simon*, 235–267.

Fenn, K. M., & Hambrick, D. Z. (2013). What drives sleep-dependent memory consolidation: greater gain or less loss? *Psychonomic bulletin & review*, *20*, 501–6.

Foer, J. (2011). *Moonwalking with Einstein: The art and science of remembering everything*. Penguin.

Fogel, S. M., Nader, R., Cote, K. a, & Smith, C. T. (2007). Sleep spindles and learning potential. *Behavioral neuroscience*, *121*, 1–10.

Fogel, S. M., & Smith, C. T. (2011). The function of the sleep spindle: a physiological index of intelligence and a mechanism for sleep-dependent memory consolidation. *Neuroscience and biobehavioral reviews*, *35*, 1154–65.

Fowler, M. J., Sullivan, M. J., & Ekstrand, B. R. (1973). Sleep and memory. *Science*.

Fritz, C. O., & Morris, P. E. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Applied Cognitive …*, *526*, 499–526.

Fuentemilla, L., Miró, J., Ripollés, P., Vilà-Balló, A., Juncadella, M., Castañer, S., … Rodríguez-Fornells, A. (2013). Hippocampus-Dependent Strengthening of Targeted Memories via Reactivation during Sleep in Humans. *Current biology : CB*, 1769–1775.

Gais, S., Mölle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *22*, 6830–4.

Genzel, L., Dresler, M., Wehrle, R., Grözinger, M., & Steiger, A. (2009). Slow wave sleep and REM sleep awakenings do not affect sleep dependent memory consolidation. *Sleep*, *32*, 302–10.

Genzel, L., Kiefer, T., Renner, L., Wehrle, R., Kluge, M., Grözinger, M., … Dresler, M. (2012). Sex and modulatory menstrual cycle effects on sleep related memory consolidation. *Psychoneuroendocrinology*, *37*, 987–98.

Geschwind, N. (1970). The Organization of Language and the Brain. *Science*, *170*, 940–944.

Gibson, B. S., Gondoli, D. M., Johnson, A. C., Steeger, C. M., & Morrissey, R. a. (2012). The future promise of Cogmed working memory training. *Journal of Applied Research in Memory and Cognition*, *1*, 214–216.

Green, C. T., Long, D. L., Green, D., Iosif, A.-M., Dixon, J. F., Miller, M. R., … Schweitzer, J. B. (2012). Will working memory training generalize to improve off-task behavior in children with attention-deficit/hyperactivity disorder? *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics*, *9*, 639–48.

Greicius, M. D., & Menon, V. (2004). Default-mode activity during a passive sensory task: uncoupled from deactivation but impacting activation. *Journal of cognitive neuroscience*, *16*, 1484–92.

Gross, A. L., & Rebok, G. W. (2011). Memory training and strategy use in older adults: Results from the active study. *Psychology and Aging*. doi:10.1037/a0022687

Hall, J., Wilson, K., & Patterson, R. (1981). Mnemotechnics: Some limitations of the mnemonic keyword method for the study of foreign language vocabulary. *Journal of Educational …*, *73*, 345–357.

Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Brain connectivity related to working memory performance. *The Journal of neuroscience*, *26*, 13338–13343.

Hampstead, B. M., Stringer, A. Y., Stilla, R. F., Giddens, M., & Sathian, K. (2012). Mnemonic strategy training partially restores hippocampal activity in patients with mild cognitive impairment. *Hippocampus*, *22*, 1652–8.

Harris, J. E. (1980). Memory aids people use: Two interview studies. *Memory & Cognition*, *8*, 31–38.

Hauck, P. D., Walsh, C. C., & Kroll, N. E. (1976). Visual imagery mnemonics: Common vs. bizarre mental images. *Bulletin of the Psychonomic Society*, *7*, 160–162.

Herrmann, D. J. (1987). TASK APPROPRIATENESS OF MNEMONIC TECHNIQUES. *Perceptual and Motor Skills*, *64*, 171–178.

Hertzog, C., & Dunlosky, J. (2012). Metacognitive approaches can promote transfer of training: Comment on McDaniel and Bugg. *Journal of Applied Research in Memory and Cognition*, *1*, 61–63.

Higbee, K. L. (1979). Recent research on visual mnemonics: Historical roots and educational fruits. *Review of Educational Research*, *49*, 611–29.

Higbee, K. L. (1997). Novices, Apprentices, and Mnemonists: Acquiring Expertise with the Phonetic Mnemonic. *Applied Cognitive Psychology*, *11*, 147–161.

Higbee, K. L. (2001). *Your memory: How it works and how to improve it*. Marlowe & Company.

Hill, R. D., Allen, C., & McWhorter, P. (1991). Stories as a mnemonic aid for older learners. *Psychology and aging*, *6*, 484–6.

Hill, R. D., Campbell, B. W., Foxley, D., & Lindsay, S. (1997). Effectiveness of the number-consonant mnemonic for retention of numeric material in community-dwelling older adults. *Experimental aging research*, *23*, 275–86.

Hossiep, R., Hasella, M., & Turck, D. (2001). *BOMAT-advanced-short version: Bochumer Matrizentest*. Hogrefe, Verlag für Psychologie.

Hrees, R. A. (1986). An edited history of mnemonics from antiquity to 1985: Establishing a foundation for mnemonic-based pedagogy with particular emphasis on mathematics. Indiana University.

Hu, Y., & Ericsson, K. A. (2012). Memorization and recall of very long lists accounted for within the Long-Term Working Memory framework. *Cognitive Psychology*, *64*, 235–266.

Hu, Y., Ericsson, K. A., Yang, D., & Lu, C. (2009). Superior self-paced memorization of digits in spite of a normal digit span: the structure of a memorist's skill. *Journal of experimental psychology. Learning, memory, and cognition*, *35*, 1426–42.

Intons-Peterson, M. J., & Fournier, J. (1986). External and internal memory aids: When and how often do we use them? *Journal of Experimental Psychology: General*, *115*, 267.

Jaeggi, S. M., Buschkuehl, M., & Etienne, A. (2007). On how high performers keep cool brains in situations of cognitive overload. *Cognitive, Affective, & …*, *7*, 75–89.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6829–33.

Julesz, B. (1971). *Foundations of cyclopean perception.* Oxford, England: U. Chicago Press.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: deciding to practice retrieval during learning. *Journal of experimental psychology. General*, *138*, 469–86.

Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review*, *1*, 602–609.

Kliegl, R., Smith, J., & Baltes, P. B. (1989). Testing-the-limits and the study of adult age differences in cognitive plasticity of a mnemonic skill. *Developmental Psychology*, *25*, 247–256.

Kliegl, R., Smith, J., & Baltes, P. B. (1990). On the Locus and Process of Magnification of Age Differences During Mnemonic Training Adults Old Adults Pretest Posttest. *Developmental Psychology*, *26*, 894–904.

Kliegl, R., Smith, J., Heckhausen, J., & Baltes, P. B. (1987). Mnemonic Training for the Acquisition of Skilled Digit Memory. *Cognition and Instruction*, *4*, 203–223.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in cognitive sciences*, *14*, 317–24.

Kondo, Y., Suzuki, M., Mugikura, S., Abe, N., Takahashi, S., Iijima, T., & Fujii, T. (2005). Changes in brain activation associated with use of a memory strategy : a functional MRI study. *NeuroImage*, *24*, 1154 – 1163.

Konrad, B. N. (2013). *Superhirn*. Wien: Goldegg Verlag GmbH.

Konrad, B. N., & Dresler, M. (2007). Außergewöhnliche Gedächtnisleistungen und Mnemotechnik. In M. Dresler (Ed.), *Wissenschaft an den Grenzen des Verstandes: Beiträge aus den Natur- und Lebenswissenschaften* (pp. 123–136). Hirzel, Stuttgart.

Konrad, B. N., & Dresler, M. (2010). Grenzen menschlicher Gedächtnisleistungen. In T. G. Baudson, A. Seemüller, & M. Dresler (Eds.), *Grenzen unseres Geistes* (pp. 65–76). Hirzel, Stuttgart.

Kroll, N. E., Schepeler, E. M., & Angin, K. T. (1986). McMurtie_Mnemonics_Educational_Applications.pdf. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 42–53.

Lahl, O., Wispel, C., Willigens, B., & Pietrowsky, R. (2008). An ultra short episode of sleep is sufficient to promote declarative memory performance. *Journal of sleep research*, *17*, 3–10.

Lawson, M. J., & Hogben, D. (1998). Learning and recall of foreign-language vocabulary: Effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction*, *8*, 179–194.

Legge, E. L. G., Madan, C. R., Ng, E. T., & Caplan, J. B. (2012). Building a memory palace in minutes: equivalent memory performance using virtual versus conventional environments with the Method of Loci. *Acta psychologica*, *141*, 380–90.

Lehrl, S., Triebig, G., & Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurologica Scandinavica*, *91*, 335–345.

Leport, A. K. R., Mattfeld, A. T., Dickinson-Anson, H., Fallon, J. H., Stark, C. E. L., Kruggel, F., … Mcgaugh, J. L. (2012). Behavioral and neuroanatomical investigation of Highly Superior Autobiographical Memory (HSAM). *Neurobiology of learning and memory*, *98*, 78–92.

Lieury, A., & Herbst, G. (2013). *Ein Gedächtnis wie ein Elefant?: Tipps und Tricks gegen das Vergessen*. Springer DE.

Louie, K., & Wilson, M. a. (2001). Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*, *29*, 145–56.

Lövdén, M., Brehmer, Y., Li, S., & Lindenberger, U. (2012). Training-induced compensation versus magnification of individual differences in memory performance. *Frontiers in human neuroscience*, *6*. doi:10.3389/fnhum.2012.00141

Luria, A. R. (1968). The mind of a mnemonist: A little book about a vast memory (L. Solotaroff, Trans.). New York: Basic Books.

Maestú, F., Simos, P. G., Campo, P., Fernández, A., Amo, C., Paul, N., … Ortiz, T. (2003). Modulation of brain magnetic activity by different verbal learning strategies. *NeuroImage*, *20*, 1110–21.

Maguire, E. A., Frackowiak, R. S. J., & Frith, C. D. (1997). Recalling routes around london: activation of the right hippocampus in taxi drivers. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *17*, 7103–10.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 4398–403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus*, *13*, 250–9.

Maguire, E. A., Valentine, E. R., Wilding, J. M., & Kapur, N. (2003). Routes to remembering: the brains behind superior memory. *Nature Neuroscience*, *6*, 90–95.

Maquet, P. (2001). The role of sleep in learning and memory. *Science (New York, N.Y.)*, *294*, 1048–52.

Martınez, A. M., & Benavente, R. (1998). The AR face database. *Rapport technique*, *24*.

Massen, C., & Vaterrodt-Plünnecke, B. (2006). The role of proactive interference in mnemonic techniques. *Memory (Hove, England)*, *14*, 189–96.

Massen, C., Vaterrodt-Plünnecke, B., Krings, L., & Hilbig, B. E. (2009). Effects of instruction on learners' ability to generate an effective pathway in the method of loci. *Memory (Hove, England)*, *17*, 724–31.

Mastropieri, M. a., Scruggs, T. E., & Levin, J. R. (1985). Mnemonic Strategy Instruction with Learning Disabled Adolescents. *Journal of Learning Disabilities*, *18*, 94–100.

McCarty, D. (1980). Investigation of a visual imagery mnemonic device for acquiring face–name associations. *… of Experimental Psychology: Human Learning and …*, *6*, 145–55.

Mcdaniel, M. A., & Bugg, J. (2012). Memory training interventions: What has been forgotten? *Journal of applied research in memory and …*, *1*, 45–50.

Mcdaniel, M. A., & Einstein, G. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *… Psychology: Learning, Memory, and …*, *12*, 54–65.

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery*.

McKenzie, G. R., & Sawyer, J. (1986). Effects of test-like practice and mnemonics on learning geographic facts. *Theory & Research in Social Education*, *14*, 201–209.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & cognition*, *29*, 10–7.

McWeeny, K. H., Young, A. W., Hay, D. C., & Ellis, A. W. (1987). Putting names to faces. *British Journal of Psychology*, *78*, 143–149.

Melby-Lervåg, M., & Hulme, C. (2013a). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*, 270–91.

Melby-Lervåg, M., & Hulme, C. (2013b). Is working memory training effective? A meta-analytic review. *Developmental psychology*, *49*, 270–91.

Merritt, J. O. (1979). None in a million: Results of mass screening for eidetic ability using objective tests published in newspapers and magazines. *Behavioral and Brain Sciences*, *2*, 612.

Merry, R. (1980). IMAGE BIZARRENESS IN INCIDENTAL LEARNING. *Psychological Reports*, *46*, 427–430.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two. *Psychological Review*, *63*, 81–97.

Moe, A., & De Beni, R. (2005). Stressing the efficacy of the Loci method: oral presentation and the subject-generation of the Loci pathway with expository passages. *Applied Cognitive Psychology*, *19*, 95–106.

Moore, C. D., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*, 11187–96.

Morris, P. E., Fritz, C. O., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for learning proper names: expanding retrieval practice, meaning and imagery. *Applied Cognitive Psychology*, *19*, 779–798.

Morris, P. E., & Greer, P. J. (1984). The effectiveness of the phonetic mnemonic system. *Human Learning: Journal of Practical Research & Applications*, *3*, 137–142.

Morris, P. E., Jones, S., & Hampson, P. (1978). An imagery mnemonic for the learning of people's names. *British Journal of Psychology*, *69*, 335–336.

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work ? The promise and challenges of enhancing cognition by training working memory, 46–60.

Nyberg, L., Sandblom, J., Jones, S., Neely, A. S., Petersson, K. M., Ingvar, M., & Bäckman, L. (2003). Neural correlates of training-related memory improvement in adulthood and aging. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 13728–33.

O'Brien, E. J., & Wolford, C. R. (1982). Effect of delay in testing on retention of plausible versus bizarre mental images. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 148–152.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature neuroscience*, *7*, 75–9.

Oswald, W. D., & Roth, E. (1987). ZVT: Zahlenverbindungstest. *Göttingen: Hogrefe*.

Ott, C. E., Butler, D. C., Blake, R. S., & Ball, J. P. (1973). The effect of interactive-image elaboration on the acquisition of foreign language vocabulary. *Language Learning*, *23*, 197–206.

Oudiette, D., Antony, J. W., Creery, J. D., & Paller, K. a. (2013). The role of memory reactivation during wakefulness and sleep in determining which memories endure. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *33*, 6672–8.

Oudiette, D., & Constantinescu, I. (2011). Evidence for the re-enactment of a recently learned behavior during sleepwalking. *PloS one*, *6*, 1–8.

Oudiette, D., & Paller, K. A. (2012). Upgrading the sleeping brain with targeted memory reactivation. *Trends in Cognitive Sciences*, 1–8.

Owen, A. M., Hampshire, A., Grahn, J. a, Stenton, R., Dajani, S., Burns, A. S., … Ballard, C. G. (2010). Putting brain training to the test. *Nature*, *465*, 775–8.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*, 46–59.

Park, D. C., Smith, a D., & Cavanaugh, J. C. (1990). Metamemories of memory researchers. *Memory & cognition*, *18*, 321–7.

Parker, E. S., Cahill, L., & McGaugh, J. L. (2006). A case of unusual autobiographical remembering. *Neurocase*, *12*, 35–49.

Patton, G. W. R. (1986). The effect of the phonetic mnemonic system on memory for numeric material. *Human Learning: Journal of Practical Research & Applications*, *5*, 21–28.

Patton, G. W. R. (1994). Testing the efficacy of name mnemonics used during conversation. *Psychological reports*, *75*, 131–142.

Patton, G. W. R., & Lantzy, P. (1987). Testing the limits of the phonetic mnemonic system. *Applied Cognitive Psychology*, *1*, 263–271.

Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., … Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron*, *44*, 535–45.

Perrin, F., García-Larrea, L., Mauguière, F., & Bastuji, H. (1999). A differential brain response to the subject's own name persists during sleep. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *110*, 2153–64.

Petrides, M., Alivisatos, B., Meyer, E., & Evans, a C. (1993). Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *90*, 878–82.

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, *9*, 534–547.

Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature neuroscience*, *3*, 85–90.

Pressley, M., Levin, J. R., & Delaney, H. (1982). The mnemonic keyword method. *Review of Educational …*, *52*, 61–91.

Rapp, S., Brenes, G., & Marsh, A. P. (2002). Memory enhancement training for older adults with mild cognitive impairment: a preliminary study. *Aging & mental health*, *6*, 5–11.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological reviews*, *93*, 681–766.

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science (New York, N.Y.)*, *315*, 1426–9.

Rasch, B., Pommer, J., Diekelmann, S., & Born, J. (2009). Pharmacological REM sleep suppression paradoxically improves rather than impairs skill memory. *Nature neuroscience*, *12*, 396–7.

Raz, A., Packard, M. G., Alexander, G. M., Buhle, J. T., Zhu, H., Yu, S., & Peterson, B. S. (2009). A slice of pi : an exploratory neuroimaging study of digit encoding and retrieval in a superior memorist. *Neurocase*, *15*, 361–72.

Rechtschaffen, A., & Kales, A. (1968). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of experimental psychology. General*, *142*, 359–79.

Richman, H. B., Staszewski, J. J., & Simon, H. a. (1995). Simulation of expert memory using EPAM IV. *Psychological review*, *102*, 305–30.

Richmond, A. S., Carney, R. N., & Levin, J. R. (2011). Got neurons? Teaching neuroscience mnemonically promotes retention and higher-order thinking. *Psychology Learning & Teaching*, *10*, 40–45.

Roediger, H. L. (1980). The Effectiveness of Four Mnemonics in Ordering Recall. *Journal of experimental psychology Human Learning and Memory*, *6*, 558–567.

Roediger, H. L., & Karpicke, J. D. (2006). The Power of Testing Memory . *Perspectives on Psychological Science*, *1*, 181–210.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *… Psychology: Learning, Memory, and …*, *21*, 803–814.

Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. a. (2009). Strengthening individual memories by reactivating them during sleep. *Science (New York, N.Y.)*, *326*, 1079.

Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage*, *9*, 216–26.

Schabus, M. (2009). Still missing some significant ingredients. *Sleep*, *32*, 291–3.

Schabus, M., Gruber, G., Parapatics, S., Sauter, C., Klösch, G., Anderer, P., … Zeitlhofer, J. (2004). Sleep spindles and their significance for declarative memory consolidation. *Sleep*, *27*, 1479–85.

Schabus, M., Hoedlmoser, K., Pecherstorfer, T., Anderer, P., Gruber, G., Parapatics, S., … Zeitlhofer, J. (2008). Interindividual sleep spindle differences and their relation to learning-related enhancements. *Brain research*, *1191*, 127–35.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is It Really Robust? *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *6*, 147–151.

Schmitt, M., Beckmann, M., Dusi, D., Maes, J., Schiller, A., & Schonauer, K. (2003). Messgüte des vereinfachten Beck-Depressions-Inventars (BDI-V). *Diagnostica*, *49*, 147–156.

Schönauer, M., Geisler, T., & Gais, S. (2013). Strengthening Procedural Memories by Reactivation in Sleep. *Journal of cognitive neuroscience*, 1–11.

Scruggs, T. E., Mastropieri, M. a., Berkeley, S. L., & Marshak, L. (2010). Mnemonic Strategies: Evidence-Based Practice and Practice-Based Evidence. *Intervention in School and Clinic*, *46*, 79–86.

Scruggs, T. E., Mastropieri, M. A., Jorgensen, C., & Monson, J. (1986). Effective mnemonic strategies for gifted learners. *Journal for the Education of the Gifted*, *9*, 105–121.

Shipstead, Z., Hicks, K. L., & Engle, R. W. (2012). Working memory training remains a work in progress. *Journal of Applied Research in Memory and Cognition*, *1*, 217–219.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological bulletin*, *138*, 628–54.

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Science (New York, N.Y.)*, *271*, 1870–3.

Smith, C. (2001). Sleep states and memory processes in humans: procedural versus declarative memory systems. *Sleep medicine reviews*, *5*, 491–506.

Soler, María JoseRulz, J. C. (1996). The Spontaneous Use of Memory Aids at Different Educational Levels. *Applied Cognitive Psychology. Feb96*, *10*.

St Clair-Thompson, H., Stevens, R., Hunt, A., & Bolder, E. (2010). Improving children's working memory and classroom performance. *Educational Psychology*, *30*, 203–219.

Stickgold, R. (2005). Sleep-dependent memory consolidation. *Nature*, *437*, 1272–8.

Stromeyer, C. F., & Psotka, J. (1970). The detailed texture of eidetic images. *Nature*, *225*, 346–9.

Susukita, T. (1933). Untersuchung eines ausserordentlichen Gedächtnisses in Japan. / A study of extraordinary memory in Japan. *Tohoku Psychologica Folia*, *1*, 111–134.

Tabatabaei, O., & Hejazi, N. H. (2011, October 26). Gender Differences in Vocabulary Instruction Using Keyword Method (Linguistic Mnemonics). *Canadian Social Science*.

Takahashi, M., Shimizu, H., Saito, S., & Tomoyori, H. (2006). One percent ability and ninety-nine percent perspiration: a study of a Japanese memorist. *Journal of experimental psychology. Learning, memory, and cognition*, *32*, 1195–200.

Thompson, C. P., Cowan, T. M., & Frieman, J. L. (1993). *Memory Search by a Memorist*. Hillsdale, NJ: Lawrence Erlbaum Associates, Incorporated.

Thompson, C. P., Cowan, T. M., Frieman, J. L., & Mahadevan, R. (1991). Rajan: A study of a memorist. *Journal of Memory and Language*, *30*, 702–724.

Thompson, T. W., Waskom, M. L., Garel, K.-L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., … Gabrieli, J. D. E. (2013). Failure of working memory training to enhance cognition or intelligence. *PloS one*, *8*, e63614.

Tilley, A. J. (1979). Sleep learning during stage 2 and REM sleep. *Biological Psychology*, *9*, 155–161.

Tombaugh, T. N. (2004). Trail Making Test A and B: normative data stratified by age and education. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, *19*, 203–14.

Troyer, A. K., Murphy, K. J., Anderson, N. D., Moscovitch, M., & Craik, F. I. M. (2008). Changing everyday memory behaviour in amnestic mild cognitive impairment: a randomised controlled trial. *Neuropsychological rehabilitation*, *18*, 65–88.

Tucker, M. a, Hirota, Y., Wamsley, E. J., Lau, H., Chaklader, A., & Fishbein, W. (2006). A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. *Neurobiology of learning and memory*, *86*, 241–7.

Turley-Ames, K. (2003). Strategy training and working memory task performance. *Journal of Memory and Language*, *49*, 446–468.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*, 273–89.

Valentine, E. R., & Wilding, J. M. (1997). *Superior Memory*. Hove, East Sussex: Psychology Press.

Valenzuela, M. J., Jones, M., Wen, W., Rae, C., Graham, S., Shnier, R., & Sachdev, P. (2003). Memory training alters hippocampal neurochemistry in healthy elderly. *Neuroreport*, *14*, 1333–7.

Van Dongen, E. V, Takashima, A., Barth, M., Zapp, J., Schad, L. R., Paller, K. A., & Fernández, G. (2012). Memory stabilization with targeted reactivation during human slow-wave sleep. *Proceedings of the National Academy of Sciences*, *109*, 10575–10580.

Vanrullen, R. (2011). Four common conceptual fallacies in mapping the time course of recognition. *Frontiers in psychology*, *2*, 365.

Vassalli, A., & Dijk, D.-J. (2009). Sleep function: current questions and new approaches. *The European journal of neuroscience*, *29*, 1830–41.

Verhaeghen, P., & Marcoen, a. (1996). On the mechanisms of plasticity in young and older adults after instruction in the method of loci: evidence for an amplification model. *Psychology and aging*, *11*, 164–78.

Voigt, U. (2001). *Esels Welt: Mnemotechnik zwischen Simonides und Harry Lorayne*. Likanas.

Wagner, A. D. (1998). Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. *Science*, *281*, 1188–1191.

Wang, A. Y., & Thomas, M. H. (1995). Effects of keyword on long-term retention: Help or hindrance? *Journal of Educational Psychology*, *87*, 468–475.

Wang, A. Y., & Thomas, M. H. (2000). Looking for long-term mnemonic effects on serial recall: The legacy of Simonides. *The American journal of psychology*, *113*, 331–340.

Weiß, R. H. (2006). Grundintelligenztest Skala 2 (CFT 20-R) mit Wortschatz-test (WS) und Zahlenfolgentest (ZF). *Intellektuelle Hochbegabung*, 80.

Wenger, M. J., & Payne, D. G. (1995). On the acquisition of mnemonic skill: Application of skilled memory theory. *Journal of Experimental Psychology Applied*, *1*, 194–215.

Wilding, J. M., & Valentine, E. R. (1988). Searching for superior memories. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. I., pp. 472–477). Wiley, Chister, UK.

Wilding, J. M., & Valentine, E. R. (1994). Memory champions. (World Memory Championships). *British Journal of Psychology*, *85*, 231–244.

Wilhelm, I., Diekelmann, S., Molzow, I., Ayoub, A., Mölle, M., & Born, J. (2011). Sleep selectively enhances memory expected to be of future relevance. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *31*, 1563–9.

Wilson, M. a, & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, *5*, 5–8.

Wollen, K. A., Weber, A., & Lowry, D. H. (1972). Bizarreness versus interaction of mental images as determinants of learning. *Cognitive Psychology*, *3*, 518–523.

World Memory Sports Council. (2013). World Memory Statistics. Retrieved August 27, 2013, from http://world-memory-statistics.com/

Worthen, J. B., & Hunt, R. R. (2011). *Mnemonology: Mnemonics for the 21st century*. Psychology Press.

Wyra, M., Lawson, M. J., & Hungi, N. (2007). The mnemonic keyword method: The effects of bidirectional retrieval training and of ability to image on foreign language vocabulary recall. *Learning and Instruction*, *17*, 360–371.

Yaakub, H., & Bakir, M. (2007). Teaching Arabic as a second language: An evaluation of key word method effectiveness. *Jurnal Teknologi*, *46*, 61–71.

Yates, F. A. (1966). *The art of memory* (Vol. 64). Random House.

Yesavage, J., & Rose, T. (1984). The effects of a face-name mnemonic in young, middle-aged, and elderly adults. *Experimental aging research*, *10*, 54–57.

Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restorative neurology and neuroscience*, *27*, 455–71.

Zelinski, E. M. (2012). Are strategies necessary to improve memory? *Journal of Applied Research in Memory and Cognition*, *1*, 56–57.