

Aus dem Max-Planck-Institut für Psychiatrie München
Direktor und Chefarzt: Prof. Dr. Martin Keck

Wirkung von Nachmittagsschlaf und wiederholtem Abruf auf die
verbale Gedächtniskonsolidierung

Dissertation

Zum Erwerb des Doktorgrades der Medizin
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von

Johannes Stintzing

aus

Starnberg

2017

Mit Genehmigung der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

Berichterstatter: Prof. Dr. med. Axel Steiger

Mitberichterstatter: Prof. Dr. Christoph J. Lauer
Prof. Dr. Till Roenneberg

Mitbetreuung durch den
promovierten Mitarbeiter: Dr. Martin Dresler

Dekan: Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung: 18.05.2017

Inhaltsverzeichnis

Abkürzungsverzeichnis	5
Begriffserklärungen	6
1 Einleitung	7
1.1 Gedächtnis	8
1.1.1 Gedächtnissysteme.....	8
1.1.2 Gedächtnisprozesse und Konsolidierung.....	9
1.2 Schlaf	10
1.2.1 Schlafstadien	11
1.2.2 Zirkadiane Rhythmik und Schlafregulation.....	12
1.2.3 Nap - Kurzschlaf	14
1.3 Schlaf und Gedächtnis	16
1.3.1 Schlaf und deklarative Gedächtniskonsolidierung.....	16
1.3.2 Neuronale Grundlagen der Gedächtniskonsolidierung im Schlaf	18
1.3.3 Schlafspindeln.....	21
1.3.4 Nap und Gedächtnis	23
1.4 Testing Effekt	25
1.4.1 Stand der Forschung	25
1.4.2 Erklärungsansätze	28
2 Zielsetzung der Arbeit	31
3 Material und Methoden	33
3.1 Studienprobanden	33
3.2 Studiendesign.....	34
3.2.1 Weitere Fragestellungen	34
3.2.2 Ablauf der Studientage	34
3.2.2.1 Ablauf der Nap- und Wachbedingung	35
3.2.2.2 Ablauf der Kontrollnap-Bedingung	37
3.3 Lernaufgabe	37
3.3.1 Entwicklung, Validierung und Programmierung der Lernaufgabe	37
3.3.2 Ablauf der Lernaufgabe	38
3.4 Polysomnographie	40
3.4.1 Schlafableitung	40
3.4.2 Auswertung der Schlafstadien und Spindeln	41
3.5 Statistik	41
4 Ergebnisse	43
4.1 Schlafauswertung.....	43

4.1.1	Strukturparameter	43
4.1.2	Spindelanalyse	44
4.2	Testauswertung	45
4.2.1	SSS und PVT	45
4.2.2	IQ-Test	46
4.2.3	Lernaufgabe	46
4.3	Inferentielle Statistik.....	51
4.3.1	Schlafvariablen und Gedächtnisleistung.....	51
4.3.2	Effekte von Schlaf und wiederholtem Testen	51
5	Diskussion	53
5.1	Zusammenfassung der Ergebnisse.....	53
5.2	Diskussion der Ergebnisse	53
5.2.1	Effekte von Schlaf und wiederholtem Testen auf die Erinnerungsleistung	53
5.2.2	Interaktion von Schlaf und Testing Effekt sowie alternative Erklärungsmodelle	55
5.2.3	Schlafstadien, Spindeln und Lernen	56
5.3	Ausblick.....	58
5.4	Diskussion der Methoden und Durchführung.....	58
6	Zusammenfassung	61
	Literaturverzeichnis	63
	Anhang	71
	Danksagung.....	72

Abkürzungsverzeichnis

AASM	American Academy of Sleep Medicine
ACh	Acetylcholin (Neurotransmitter)
ANOVA	Varianzanalyse
ARAS	aufsteigendes retikuläres Aktivierungssystem
CWT	Continuous Wavelet Transform
EEG	Elektroenzephalographie
EOG	Elektrookulographie
EMG	Elektromyographie
fMRT	funktionelle Kernspintomographie
GABA	Gamma-Hydroxy-Buttersäure (Neurotransmitter)
M	Mittelwert
MANOVA	Multivariate Varianzanalyse
MWT	Mehrfach-Wortschatz-Intelligenz-Test
N1,N2,N3	Schlafstadien 1-3
NREM	Non-REM Schlafstadien
PET	Positronen-Emissions-Tomographie
PVT	Psychomotor Vigilance Task
REM	Rapid Eye Movement (Schlafstadium)
SCN	Nucleus Suprachiasmaticus
SSS	Stanford Sleepiness Scale
St.Abw.	Standardabweichung
SWS	Slow Wave Sleep, entspricht Stadium N3
T1,T2, T3	beschreibt im Rahmen dieser Arbeit die Tests im zeitlichen Verlauf
VLPO	Nucl. preopticus ventrolateralis

Begriffserklärungen

Baseline	Im Rahmen dieser Arbeit erster Test direkt nach dem Üben
Item	einzelnes Element einer größeren Reihe/Gruppe
Nap	kurzer (Nachmittags-) Schlaf
Retest	Im Rahmen dieser Arbeit der online-basierte Vokabeltest nach vier Tagen
Testing	wiederholter Abruf des Gelernten aus dem Gedächtnis

1 Einleitung

Seit Herrmann Ebbinghaus 1885 die Ergebnisse seiner Selbstversuche veröffentlichte, beschäftigt sich die Wissenschaft mit dem Vergessen. In seinen Versuchen hatte er sinnlose Silbenreihen auswendig gelernt und dann in verschiedenen Abständen überprüft, wie stark das Vergessen fortgeschritten war (Ebbinghaus 1885). Die daraus entwickelte Vergessenskurve zeigt graphisch, was auch die Erfahrung lehrt, nämlich dass ein großer Teil dessen, was wir in Schule, Studium oder Berufsleben erlernen, bei unzureichender Anwendung oder Wiederholung schon bald wieder aus dem Gedächtnis verschwunden sein wird. Diese Vergänglichkeit der Erinnerung wird von Daniel Schacter als eine der „Sieben Sünden des Gedächtnisses“ bezeichnet (Schacter 2001). Doch erfolgreiches Lernen gehört zu den wichtigsten Voraussetzungen für Erfolg in unserer Bildungsgesellschaft, in der aufgrund der rasanten Fortschritte in Wissenschaft und Technik lebenslanges Lernen für die meisten Menschen unabdingbar geworden ist. Wie Wissen erfolgreich erworben und behalten werden kann, ist einer der Hauptforschungsbereiche in Psychologie und Pädagogik sowie Thema zahlreicher Ratgeber. Auch bedienen sich nicht wenige Studenten pharmakologischer Mittel (sog. Hirndoping), um ihre Gedächtnisleistung in Prüfungszeiten zu verbessern. Die vorliegende Arbeit beschäftigt sich mit der Erforschung zweier Verhalten, die nachgewiesenermaßen einen positiven Effekt auf das langfristige Behalten von erlerntem Wissen haben: Das wiederholte Abtesten des Gelernten, genannt „Testing Effekt“, und Schlaf, sowie deren mögliche Interaktion untereinander.

1.1 Gedächtnis

Die Fähigkeit des Menschen, Wissen und Erfahrungen zu sammeln, im Gedächtnis zu speichern und bei Bedarf abzurufen, ist eine Grundvoraussetzung für die persönliche Entwicklung und die erfolgreiche Anpassung an die verschiedenartigen Anforderungen des Lebens (Rasch und Born 2013).

1.1.1 Gedächtnissysteme

Es werden verschiedene Formen des Gedächtnisses unter zeitlichen und qualitativen Aspekten unterschieden¹: Mit Blick auf die Verweildauer der Gedächtnisinhalte wird das Gedächtnis in das sensorische Gedächtnis, das Kurzzeitgedächtnis, und das Langzeitgedächtnis unterteilt. Im *sensorischen Gedächtnis* werden Sinneseindrücke für die Dauer von wenigen hundert Millisekunden festgehalten, die wichtigsten Inhalte extrahiert und gespeichert, während der Großteil direkt wieder verloren geht. Das *Kurzzeitgedächtnis* kann nur eine sehr begrenzte Anzahl an verbal codierten Items für die Dauer von einigen Sekunden bis maximal Minuten speichern. Ein klassisches Beispiel hierfür wäre die Telefonnummer, die man sich für die Zeit des Wählens merkt, aber kurz danach wieder vergessen hat. Um Gedächtnisinhalte dauerhaft zu speichern, müssen diese ins *Langzeitgedächtnis* überführt werden, das eine nahezu unbegrenzte Kapazität hat.

Inhaltlich wird das Langzeitgedächtnis unterteilt in das deklarative und das nicht-deklarative Gedächtnis: das *deklarative Gedächtnis* enthält episodisches und semantisches Wissen. Episodisches Gedächtnis bezeichnet raum-zeitlich eingebettete Ereignisse sowie auch autobiografisches Wissen. Semantisches Wissen bezeichnet kontextunabhängiges Faktenwissen. Diese Gedächtnisinhalte sind durch aktive Erinnerung zugänglich (Rasch und Born 2013). Für die Bildung deklarativer Gedächtnisinhalte sind der Hippocampus und angrenzende Areale von elementarer Bedeutung. Eine bilaterale Läsion des Hippocampus führt zu anterograder Amnesie, d.h. neue deklarative Inhalte können nicht mehr erfolgreich im Gedächtnis gespeichert werden. Die Bildung nicht-deklarativer Gedächtnisinhalte hingegen ist von der Amnesie nicht betroffen. Dazu zählen u.a. motorische Routinen, Priming sowie klassische Konditionierung (Squire und Zola 1996).

¹ Folgende Übersicht gemäß Schmidt 2007, S. 227–228; Zimbardo und Gerrig 2008, S. 236–246

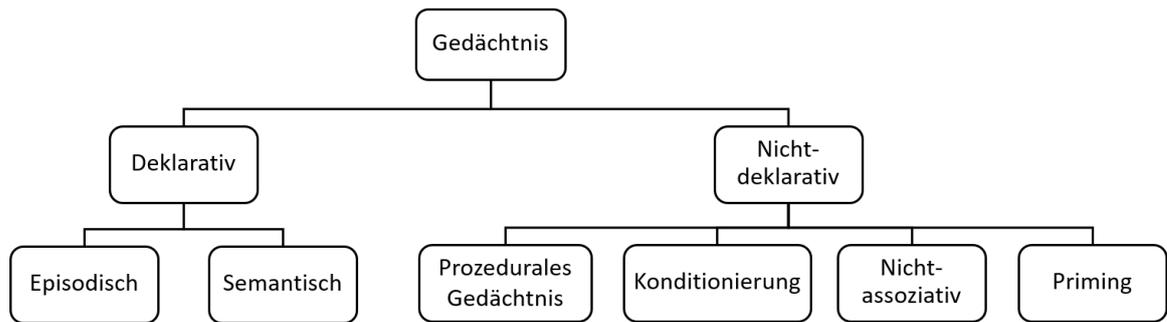


Abb. 1: Gedächtnissysteme; modifiziert nach Walker und Stickgold 2006

1.1.2 Gedächtnisprozesse und Konsolidierung

Man kann die Gedächtnisprozesse in drei Phasen einteilen: während der *Enkodierung* wird zu einer Wahrnehmung eine Repräsentation im Gedächtnis gebildet, eine Gedächtnisspur. Diese ist zu Beginn noch relativ störanfällig, wird aber im nachfolgenden Prozess der *Konsolidierung* verstärkt und in existierendes Wissen integriert. Beim *Abruf* wird auf dieses Wissen zugegriffen und es wieder ins Bewusstsein gerufen (Rasch und Born 2013). Die *Konsolidierungstheorie* nach Müller und Pilzecker aus dem Jahr 1900 besagt, dass nach Abschluss des aktiven Lernens im Gehirn noch für einige Zeit physiologische Prozesse stattfinden, die die entstandenen Assoziationen konsolidieren und festigen (Lechner et al. 1999). In dieser ersten Phase der Konsolidierung kann die Festigung der Gedächtnisspur durch Interferenzlernen, elektrische Schocks, oder pharmakologische Agenzien unterbrochen oder zumindest empfindlich gestört werden (Lechner et al. 1999; McGaugh 1966; Wixted 2004). Auf neuronaler Ebene geht man von zwei Formen der Konsolidierung aus, der „Synapsenkonsolidierung“ und der „Systemkonsolidierung“. Entsprechend der Vorstellung der synaptischen Plastizität kommt es während der Konsolidierung einer Erinnerung zu biochemischen Prozessen an den betroffenen Synapsen, die nachfolgend zu einem Umbau der Neuronenfortsätze und damit auch zu einer effizienteren Kommunikation zwischen den beteiligten Neuronen führen. Diese Synapsenkonsolidierung geschieht während der ersten Minuten und Stunden nach dem Lernen. Die Systemkonsolidierung hingegen beschreibt einen Prozess, der über Wochen, Monate, oder sogar Jahre stattfindet. Hierbei wird die Erinnerung zunehmend unabhängig vom Hippocampus und dafür verstärkt in anderen Hirnregionen verankert. Dies erklärt, warum bei Gedächtnisverlust häufig die am kürzesten

zurückliegenden Lebensereignisse am stärksten betroffen sind. In den vergangenen Jahren wurde auch vermehrt das Konzept der Rekonsolidierung, also der erneuten Konsolidierung im Rahmen des Abrufs einer Erinnerung, in Ergänzung zu der nach dem Erlernen stattfindenden zeitlich begrenzten Konsolidierung diskutiert (Dudai 2004; Übersicht: Rasch und Born 2013)

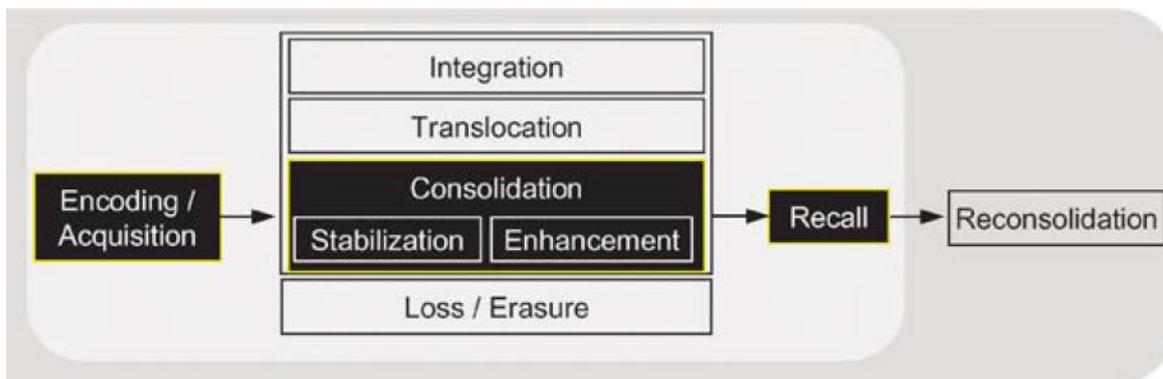


Abb. 2: Übersicht über die Gedächtnisprozesse (Walker und Stickgold 2006)

1.2 Schlaf

Der Mensch verbringt etwa ein Drittel jedes Tages mit Schlafen. Schlaf ist gekennzeichnet durch einen Bewusstseinsverlust sowie relative körperliche Inaktivität (Rasch und Born 2013). Während dieser Zeit kann keine produktive Arbeit verrichtet werden und gleichzeitig stellt der Bewusstseinsverlust mit verminderten Reaktionen auf äußere Reize ein Sicherheitsrisiko dar. Die Tatsache, dass trotzdem alle Menschen und auch viele Tiere schlafen müssen, legt die Vermutung nahe, dass Schlaf einen wichtigen Zweck erfüllt, der diese Nachteile mehr als ausgleicht (Rattenborg et al. 2007). Auch die eigene Erfahrung lehrt, dass die Leistungsfähigkeit durch Schlafmangel deutlich eingeschränkt sein kann und dass der Körper nach einer durchwachten Nacht versuchen wird, einen Teil des verpassten Schlafes nachzuholen.

Es gibt eine Reihe von Annahmen und viel Forschung bezüglich der Funktion des Schlafes: Erholung und Regeneration, verminderter Energieverbrauch, Thermoregulation und Effekte auf das Immunsystem (Stuck und Fricke-Oerkermann 2013). Dafür wäre aber möglicherweise auch ein Zustand körperlicher Ruhe ohne Bewusstseinsverlust und Abschirmung von äußeren Reizen ausreichend, so dass die Vermutung naheliegt, dass Schlaf vor allem für das Gehirn eine wichtige Funktion erfüllt (Rasch und Born 2013). Schlafmangel hat primär negative Folgen für die kognitive Leistungsfähigkeit, die über

den Effekt der eingeschränkten Vigilanz und Aufmerksamkeit hinausgehen (Killgore 2010).

1.2.1 Schlafstadien

Schlaf ist kein einheitlicher Zustand, sondern setzt sich aus verschiedenen Schlafstadien zusammen. Diese werden anhand charakteristischer EEG-Merkmale (Amplitude und Frequenz), dem gemessenen Muskeltonus und Augenbewegungen eingeteilt. Die ältere Einteilung nach Rechtschaffen und Kales (1968) unterschied die Stadien S1-S4 und REM (s. Abb. 3). Seit 2007 gibt es die neue Klassifikation der AASM (Iber et al. 2007), die u.a. die früheren Tiefschlafstadien S3 und S4 zu einem Stadium N3 zusammenfasst. Durch Veränderungen in den Kriterien kommt es darüber hinaus zu Abweichungen bei der Zuordnung zu den einzelnen Stadien und folglich deren Länge (Moser et al. 2009). Die AASM-Klassifikation unterteilt den Schlaf in die Non-REM Stadien N1-N3 und den REM-Schlaf²: Das Stadium N1 bezeichnet den oberflächlichen Schlaf, der direkt nach dem Einschlafen auftritt und eine Art Übergang zwischen Wachzustand und Schlaf darstellt. Kennzeichnend ist eine Theta-Aktivität im EEG sowie langsame rollende Augenbewegungen. Wird ein Proband in den ersten Minuten nach dem Einschlafen wieder wach, besteht kein Bewusstsein bezüglich der Tatsache, dass er schon geschlafen hat. Das Stadium N2 wird auch als stabiler Schlaf bezeichnet und hat zusätzlich zur Theta-Aktivität auch Schlafspindeln und K-Komplexe im EEG, dafür aber kein Augenrollen. Im Stadium N3, das auch als Tiefschlaf oder Slow-Wave-Sleep (SWS) bezeichnet wird, sinkt die Frequenz und steigt die Amplitude der EEG-Kurven, man sieht Delta-Wellen³ mit einer Frequenz < 2 Hz. Der Muskeltonus sinkt mit der Schlaftiefe über die Stadien N1-N3 kontinuierlich ab. Das Stadium REM (Rapid Eye Movement) erhält seinen Namen von den charakteristischen schnellen, konjugierten Augenbewegungen. Es wird auch als paradoxer Schlaf bezeichnet, denn obwohl das EEG dem eines wachen und aktiven Gehirns ähnelt, ist die Weckschwelle in diesem Stadium am höchsten und der Muskeltonus sehr niedrig. Im EEG finden sich neben einer Theta- und langsamen Alpha-Aktivität auch Sägezahnwellen.

² Beschreibung gemäß Stuck und Fricke-Oerkermann 2013, sowie Scoring-Regeln der Amerikanischen Akademie für Schlafmedizin (AASM) (Iber et al. 2007)

³ Delta Wellen sind definiert als EEG-Rhythmus bestehend aus 1-4 Hz Aktivität, allerdings werden im Rahmen des visuellen Scoring nur Frequenzen < 2 Hz dem SWS zugeordnet (Iber et al. 2007)

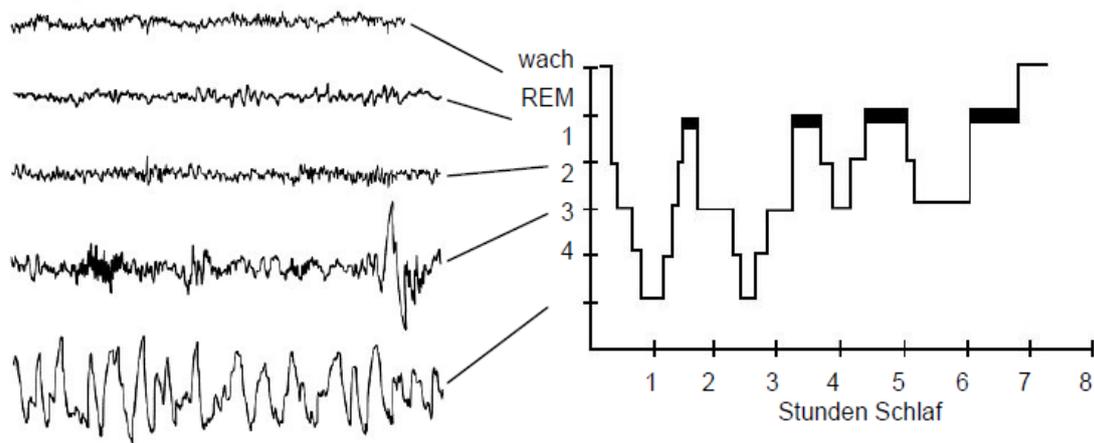


Abb. 3: Hypnogramm (rechts) mit der klassischen Schlafarchitektur aus mehreren Zyklen. In der ersten Nachthälfte mehr Tiefschlaf, in der zweiten Hälfte mehr REM. Links die entsprechenden typischen EEG-Ableitungen der unterschiedlichen Schlafstadien (Dresler 2007).

Nach dem Einschlafen werden die verschiedenen Stadien typischerweise in der Reihenfolge N1 - N2 - N3 - REM durchlaufen, dies bezeichnet man als Schlafzyklus. Diese Zyklen wiederholen sich etwa 4- bis 6-mal pro Nacht, wobei jeder einzelne Zyklus zwischen 90-110 Minuten dauert (Stuck und Fricke-Oerkermann 2013). Die Charakteristika des Schlafes, wie durchschnittliche Schlafzeit pro Tag, Verteilung des Schlafes (zusammenhängend oder fragmentiert) und die Schlafarchitektur (Anteil der Schlafstadien am Gesamtschlaf), variiert immens in Abhängigkeit vom Lebensalter (Wauquier 1993; Galland et al. 2012). Im Folgenden soll nur auf die Charakteristika bei gesunden, jungen Erwachsenen eingegangen werden. Während der ersten Nachthälfte ist der Anteil an Tiefschlaf deutlich größer als der Anteil an REM, während in der zweiten Nachthälfte mehr REM als SWS auftritt (Plihal und Born 1997).

1.2.2 Zirkadiane Rhythmik und Schlafregulation

Schlaf ist überlebenswichtig für den Menschen. Deshalb ist es sinnvoll, dass dieser wichtige Vorgang nicht von einer einzelnen Hirnregion oder physiologischem Mechanismus abhängt (Dresler et al. 2014). Stattdessen sind an der Regulation des Schlafes verschiedene Gehirnregionen, Transmittersysteme und Hormone beteiligt⁴. Auf anatomischer Ebene spielen hierbei Hypothalamus und Hirnstamm eine zentrale Rolle

⁴ Auf die beteiligten Transmittersysteme und Hormone wird im Folgenden nicht detailliert eingegangen. Für weiterführende Informationen hierzu s. Dresler et al. 2014

(Dresler et al. 2014). Ausgehend vom sog. Aufsteigenden retikulären Aktivierungssystem (ARAS) führen Projektionen in Thalamus und Kortex und bewirken Wachheit. Während des Schlafes wird das ARAS durch GABAerge Projektionen aus dem Nucl. preopticus ventrolateralis (VLPO) gehemmt. Umgekehrt hemmt das aktive ARAS im Wachzustand die Neuronen des VLPO. Durch diese gegenseitige Hemmung entsteht eine Art neuronaler „Schalter“, der zu scharfen Übergängen zwischen Schlaf- und Wachzustand führt, ohne dass der Mensch viel Zeit in Übergangsphasen verbringt (Saper et al. 2005). Auch die einzelnen Schlafphasen werden über neuronale Netzwerke gesteuert. So bestimmen glutamaterge und GABAerge Neuronen des Hirnstamms mit ihrem Wechselspiel den Zeitpunkt des Eintritts in das REM-Stadium (Luppi et al. 2012).

Ein einflussreiches und intuitiv nachvollziehbares Modell zur Beschreibung der Schlafregulation ist das Zwei-Prozess-Modell, welches Borbély 1982 formulierte: Prozess S ist homöostatisch reguliert und nimmt als Schlafdruck mit Dauer der Wachphase kontinuierlich zu und wird im Schlaf abgebaut. Prozess C beschreibt die zirkadiane Komponente der Einschlafneigung, unabhängig von der Dauer der Wachphase (Borbély und Achermann 1992; Saper et al. 2005). Als möglicher Mechanismus für Prozess S gilt die Akkumulation von Adenosin in der Wachphase, das als schlafförderndes Substrat wirkt. Der Nucleus suprachiasmaticus (SCN) des Hypothalamus fungiert als zentraler Zeitgeber für den zirkadianen Rhythmus und entspricht damit der Komponente des Prozess C (Saper et al. 2005). In Bunkerexperimenten konnte gezeigt werden, dass auch unabhängig von externer Information über Tageszeit Menschen mit einem endogenen Tag-Nacht-Rhythmus von etwa 25 h weiterleben, wobei hier interindividuelle Unterschiede bestehen (Aschoff 1965). Die innere Uhr wird primär über Sonnenlicht auf den externen 24h-Rhythmus synchronisiert (Roenneberg et al. 2007), wobei der SCN diese Informationen von spezialisierten Zellen der Retina erhält (Saper et al. 2005). In Experimenten, in denen Probanden nach einem erzwungenen 20-Stunden-Tag leben, kommt es zu einer Dissoziation und Asynchronität des zirkadianen Rhythmus und des Schlaf-Wach-Rhythmus. Hierdurch lassen sich die Effekte der zirkadianen und homöostatischen Komponente getrennt untersuchen. Es zeigt sich, dass in der zirkadianen Hochphase die kognitive Leistungsfähigkeit hoch, die Einschlafzeit deutlich verlängert, die Schlafdauer und Schlafqualität vermindert und der Schlaf insgesamt stärker durch Wachphasen unterbrochen war. In der zirkadianen Tiefphase hingegen wird Schlaf durch die zirkadiane Komponente unterstützt, wobei diese eine wichtige Rolle für

die Kontinuität der Schlafperiode spielt, insbesondere gegen Ende der Schlafperiode, wenn der homöostatische Schlafdruck deutlich geringer ist (Wyatt et al. 1999).

1.2.3 Nap - Kurzschlaf

Als Nap oder Kurzschlaf werden Schlafperioden von drei Minuten bis drei Stunden Dauer bezeichnet (Medrick und Drummond 2009). Während der Nap im Kleinkindesalter regelhaft auftritt und auch im hohen Lebensalter wieder deutlich zunimmt, gibt es bei jungen Erwachsenen große interindividuelle Unterschiede in Bezug auf die Häufigkeit und den empfundenen Nutzen des Naps (Milner und Cote 2009). Insbesondere in südlichen Regionen hat die nachmittägliche Ruhepause in Form der Siesta eine lange Tradition, aber auch in unseren Breiten erfreut sich das „Nickerchen“ oder der „Powernap“ einer gewissen Beliebtheit. Auch die Zunahme der Schichtarbeit, der internationalen Reisen und die Abnahme der durchschnittlichen nächtlichen Schlafdauer tragen zu vermehrter Tagesmüdigkeit und folglich vermehrtem Tagschlaf bei (Medrick und Drummond 2009). Je nach zugrundeliegender Motivation können verschiedene Formen des Naps unterschieden werden: der prophylaktische Nap, z.B. vor einer durchwachten Nacht, hat einen länger anhaltenden positiven Effekt auf Leistungsfähigkeit und Wachheit als der Gebrauch von Koffein (Bonnet et al. 1995). Der Nachhol-Nap dient dem Ausgleich negativer Folgen des vorangegangenen Schlafmangels, insbesondere vermehrte Schläfrigkeit (Einschlafneigung) und Erschöpfung (eingeschränkte kognitive Leistungsfähigkeit) (Medrick und Drummond 2009). Aber auch bei Menschen ohne Schlafmangel kann ein kleines Nickerchen zu einem besseren Lebensgefühl und besseren kognitiven Leistungen beitragen (Milner und Cote 2009).

Ein Nap kann prinzipiell zu jeder Tages- und Nachtzeit stattfinden, allerdings gibt es ein Zeitfenster zwischen 14 und 16 Uhr, das sog. Nachmittagstief (das auch unabhängig von einer vorhergegangenen reichlichen Mahlzeit auftritt), wo die Schlafneigung besonders hoch ist. Dies zeigt sich auch im MSLT (Multiple Sleep Latency Test), der die Einschlaflatenz zu verschiedenen Zeiten über den Tag verteilt misst und regelhaft ein Minimum während dieses Nachmittagstiefs zeigt. Als physiologisches Korrelat zeigt sich ein Abfall der Körperkerntemperatur, die über den Tag mit einem 12h bicircadianen Kurvenverlauf mit zwei Tiefpunkten - von 2-4 Uhr sowie von 14-16 Uhr – beschrieben werden kann (Medrick und Drummond 2009). Abschließend lässt sich allerdings sagen, dass der optimale Zeitpunkt für den Nap von individuellen Faktoren wie Schlafbedürfnis,

Tag/Nacht-Rhythmus, Chronotyp und vorangegangener Wachzeit abhängt und somit nicht allgemeingültig festgelegt werden kann (Milner und Cote 2009).

Auch die optimale Napdauer hängt von verschiedenen Faktoren, insbesondere dem erwünschten Effekt, ab: Ein kurzer „Powernap“ von etwa 10 Minuten Dauer führt zu einer Verminderung der Schläfrigkeit und Erschöpfung, mehr Energie und besserer kognitiver Leistungsfähigkeit direkt nach dem Nap für die folgenden 2-3 Stunden. Längere Naps führen zu ähnlichem Nutzen, allerdings erst nach einer Phase der Schlaftrunkenheit mit Abgeschlagenheit und kognitiver Verlangsamung, die möglicherweise mit der Gesamtschlafdauer oder aber der Menge an enthaltenem Tiefschlaf zusammenhängt (Brooks und Lack 2006). Eine andere Studie, die Schichtarbeit oder lang andauernde Arbeitseinsätze simuliert, mit einem Nap nach 20 bzw. 30h Wachzeit, kommt zu dem Ergebnis, dass die Leistungsfähigkeit in den darauffolgenden Stunden von einer längeren Napdauer von 40-60 Minuten mehr profitiert als von kurzen Naps (Mulrine et al. 2012). Allgemein gesprochen sind Naps sehr spannend für die Schlafforschung, da es über die Veränderung der Napzeit und -dauer gut möglich ist, zu beeinflussen, welche Schlafstadien darin enthalten sind. So enthalten sehr kurze Naps nur Stadium 1 oder Stadium 1 und 2, längere Naps enthalten auch Tiefschlaf und ggf. REM-Schlaf (Medrick und Drummond 2009). Damit lassen sich elegant methodische Schwachpunkte der Schlafstadienisolierung bei Nachtschlaf, wie durch wiederholtes Wecken oder Vergleich von Nachthälften, die immer in geringem Maße auch die anderen Stadien enthalten, vermeiden und darüber hinaus modifizierende Faktoren wie Effekte des Schlafmangels in der wachen Kontrollgruppe, zirkadiane Effekte oder Effekte der nächtlichen Hormonsekretion verringern (Ficca et al. 2010; Lahl et al. 2008; für einen Vergleich der Methoden s.a. Genzel et al. 2015b).

1.3 Schlaf und Gedächtnis

Mirum dictu est nec in promptu ratio, quantum nox interposita adferat firmitatis, sive requiescit labor ille, cuius sibi ipsa fatigatio obstabat, sive maturatur atque concoquitur, quae firmissima eius pars est, recordatio; quae statim referri non poterant, contexuntur postera die, confirmatque memoriam idem illud tempus, quod esse in causa solet oblivionis.

Quintilian⁵

Die wissenschaftliche Erforschung von Schlaf und Gedächtnis blickt auf eine nahezu hundertjährige Geschichte zurück. Schon Ebbinghaus (1885) stellte bei seinen Experimenten fest, dass es zu einer Verlangsamung des Vergessens durch Schlaf kam. Dies stellte den Ausgangspunkt für erste Studien zum Thema Schlaf und Gedächtnis dar (z.B. Jenkins und Dallenbach 1924). Seither wurden zahlreiche Studien im Bereich Schlaf und Gedächtnis durchgeführt. In ihrer Übersichtsarbeit ziehen Diekelmann und Born (2010) angesichts der breiten Literatur das folgende Fazit: „Numerous studies have confirmed the beneficial effect of sleep on declarative and procedural memory in various tasks, with practically no evidence for the opposite effect (sleep promoting forgetting).”

1.3.1 Schlaf und deklarative Gedächtniskonsolidierung

Es konnte gezeigt werden, dass nicht nur der Schlaf nach dem Lernen wichtig ist, sondern auch ausreichend Schlaf vorher ist von großer Bedeutung. Schlafentzug über 36 Stunden führte in einer Studie zu 40% schlechteren Ergebnissen in einer deklarativen Gedächtnisaufgabe, was auf eine Einschränkung der Enkodierleistung zurückzuführen ist (Walker und Stickgold 2006). Zumeist geht es in der Forschung jedoch um den Schlaf, der zwischen Enkodierung und Abruf stattfindet. Hierzu gibt es in der Literatur konkurrierende Ansichten, ob und wie die deklarativen Gedächtnisinhalte durch Schlaf beeinflusst werden. Es gibt vereinzelte Stimmen, die argumentieren, es gebe keinen

⁵ Zitat aus Quintilian`s Institutio Oratoria XI 2, 43:

„Seltsam und nicht leicht zu erklären ist die Tatsache, wieviel der Zwischenraum einer Nacht zur Befestigung des Gedächtnisses ausmacht, sei es deshalb, weil die Anstrengung, bei der die Erschöpfung ihr selbst entgegenwirkte nun ruht, sei es auch, weil nun gereift und gar wird, was der wichtigste Teil des Gedächtnisses ist, nämlich die Wiedererinnerung – jedenfalls: was auf der Stelle sich nicht einstellen wollte, läuft am folgenden Tag wie am Schnürchen und seine Befestigung verdankt das Gedächtnis gerade der Zeit, die doch das Vergessen zu verursachen pflegt.“ (Quintilian 1995, S. 605)

Zusammenhang zwischen Schlaf und Konsolidierung. Dabei wird unter anderem auf Studien mit REM-Entzug bzw. Individuen ohne REM-Schlaf (nach Schädigung des Hirnstamms oder unter antidepressiver Medikation) verwiesen, die zu keinen Einschränkungen in der allgemeinen Leistungsfähigkeit führten (Vertes 2004; Vertes und Siegel 2005). Im Hinblick auf deklarative Gedächtnisinhalte scheint allerdings NREM eine wichtigere Rolle zu spielen als REM (Plihal und Born 1997; Born et al. 2006). Insgesamt gibt es eine stetig wachsende Zahl an Studien, die einen positiven Effekt von Schlaf verglichen mit dem gleichen Zeitraum im Wachzustand belegen (u.a. Jenkins und Dallenbach 1924; Ellenbogen et al. 2006a; Plihal und Born 1997).

Es gibt verschiedene theoretische Erklärungsansätze für diesen Schlafeffekt: Die *Interferenztheorie des Vergessens* sieht Schlaf primär als einen passiven Zustand, der durch Reizabschirmung von der Umwelt neu erworbene Gedächtnisinhalte vor der Interferenz durch weitere Eindrücke schützt. Gegen eine rein passive Rolle des Schlafes spricht allerdings die Zeitabhängigkeit des Effekts, d.h. Schlaf kurz nach der Enkodierung schützt besser als Schlaf einige Stunden später bei gleicher Schlafdauer und gleichem Intervall bis zum Abruf (Gais 2006; Rasch und Born 2013). Im Rahmen einer Studie lernten Probanden Wortpaare und nach 12 sowie nach 24 Stunden wurde getestet, wie viel sie behalten hatten. Innerhalb dieser 24 Stunden schlief die Hälfte der Probanden während der ersten 12 Stunden und war während der zweiten zwölf Stunden wach, während die andere Hälfte während der ersten zwölf Stunden wach war und während der zweiten zwölf Stunden schlief. Bei dem Test nach 24 Stunden schnitt die Gruppe besser ab, die zuerst geschlafen hatte. Das lag daran, dass die Gruppe, die zuerst geschlafen hatte während ihrer Wachphase deutlich langsames Vergessen zeigte als die Gruppe, die zuerst wach war und erst danach schlief. Dies stützt die Annahme, dass die Erinnerungen im Schlaf so stabilisiert wurden, dass sie in der darauffolgenden Wachphase resistenter gegen das Vergessen waren (Payne et al. 2012). In weiteren Studien fanden sich auch qualitative Veränderungen der Erinnerung nach Schlaf, wie beispielsweise das Erkennen einer zusätzlichen zugrundeliegenden Regel, die bei der Bearbeitung von Zahlenreihen zu einem deutlich Zeitgewinn bei der Lösung führt (Wagner et al. 2004; Diekelmann und Born 2010). Diese Erkenntnisse sind vereinbar mit der *Konsolidierungshypothese*, wobei hier Schlaf entweder als eine opportunistische Phase gesehen wird, in der die normalen Konsolidierungsvorgänge aufgrund von weniger Interferenz ungestörter verlaufen können als im Wachzustand, oder aber als aktiver Prozess, der die Konsolidierung

aufgrund seiner besonderen physiologischen Vorgänge in elementarer Weise fördert (Ellenbogen et al. 2006b).

Die Konsolidierung von Gedächtnisinhalten im Schlaf wird u.a. durch den emotionalen Gehalt (Payne et al. 2008), Dauer des Schlafes, Motivation, Abfrageform und einer Reihe anderer Faktoren beeinflusst (Übersicht: Diekelmann et al. 2009). In einer Studie zeigte sich nach einem zwölfstündigen Intervall mit oder ohne Schlaf ein signifikanter Unterschied bei Wortpaaren mit unabhängigen Begriffen, nicht jedoch bei Wortpaaren mit themenverwandten Begriffen (Payne et al. 2012). Um vergleichbare Ergebnisse zu erhalten scheint es daher wichtig, gleiche oder sehr ähnliche Testmethoden zu verwenden, da Unterschiede in der Schwierigkeit oder anderen Aspekten des Testdesigns möglicherweise zu einer anderen Verarbeitung und entsprechend stark abweichenden Ergebnissen führen können.

Deklarative Gedächtniskonsolidierung korreliert positiv mit SWS. Dies zeigte u.a. eine Studie von Plihal und Born (1997), die Lernerfolg für eine deklarative Aufgabe nach der ersten (SWS-reichen) bzw. der zweiten (REM-reichen) Nachthälfte verglich. Während die sog. *Zwei-Prozess-Hypothese* davon ausgeht, dass deklarative Inhalte während SWS und prozedurale Inhalte während REM konsolidiert werden, betonen *Sequentielle Hypothesen* die Bedeutung der Abfolge von NREM und REM, die sich mehrfach zyklisch wiederholen (Piosczyk et al. 2009). So führt die Segmentierung des Schlafes durch wiederholtes Wecken mit Störung der zyklischen NREM-REM Abfolge zu einer deutlich eingeschränkten Erinnerung an gelernte Wortpaare. Dieser Effekt trat nicht auf bei segmentiertem Schlaf mit erhaltener NREM-REM Struktur bei ansonsten vergleichbarer Gesamtschlafdauer und Dauer der einzelnen Stadien (Ficca et al. 2000). In den vergangenen Jahren ist auch der Leichtschlaf, insbesondere NREM Stadium 2 in den Fokus der Gedächtnisforschungen gerückt. Während Effekte des Leichtschlafes in Studien, in denen Nachthälften verglichen werden, nicht zu isolieren sind, lässt sich anhand kurzer Naps ohne SWS und REM zeigen, dass es auch hier positive Effekte auf die Gedächtniskonsolidierung gibt (Genzel et al. 2014).

1.3.2 Neuronale Grundlagen der Gedächtniskonsolidierung im Schlaf

Wenn man die positiven Effekte von Schlaf auf Gedächtnis betrachtet, stellt sich die Frage, welche Prozesse im Gehirn ablaufen, um die Stabilisierung und Verbesserung der Erinnerung zu bewirken. Die Forschung konnte hier in den vergangenen Jahren einige

neue Erkenntnisse liefern. So zeigten Versuche mit Ratten, dass bestimmte Erregungsmuster im Hippocampus, die beim Ablaufen einer bestimmten Route stattfinden, im darauffolgenden Schlaf (bevorzugt im SWS) wiederholt reaktiviert werden. Dabei können auch andere Bereiche des Gehirns wie das Striatum und der Kortex Reaktivierungen zeigen (Wilson und McNaughton 1994; Qin et al. 1997; Lansink et al. 2009; für Übersichtsarbeiten s. Sutherland und McNaughton 2000; Rasch und Born 2013). Mithilfe neuerer bildgebender Verfahren wie PET und fMRT konnten auch beim Menschen Anzeichen für Reaktivierung gefunden werden: Nach dem Erlernen von Gesicht – Szene Assoziationen, konnten bei gleichzeitiger EEG-Ableitung und fMRT im nachfolgenden NREM-Schlaf vermehrte gemeinsame Aktivierungen des Hippocampus und der für Gesichter und Szenen spezifischen visuellen Kortexareale gezeigt werden. Diese Aktivierungen standen in zeitlichem Zusammenhang mit Schlafspindeln im EEG (Bergmann et al. 2012). In einer anderen Studie lernten Probanden die Lokalisation von Kartenpaaren (Memory), während ihnen ein Duft präsentiert wurde. Wurde ihnen derselbe Duft während des nachfolgenden SWS erneut präsentiert, zeigte sich im fMRT eine Aktivierung des Hippocampus und eine verbesserte Merkleistung nach dem Schlaf. Dieser positive Effekt trat jedoch nicht auf, wenn der Duft während Wachzustand oder REM präsentiert wurde oder wenn er beim Enkodieren nicht präsentiert worden war. Aus diesem Ergebnis schlossen die Autoren, dass die Reaktivierungen einen „kausalen Faktor“ in der Konsolidierung der Gedächtnisinhalte darstellen (Rasch et al. 2007). Möglicherweise machen gerade diese Reaktivierungen und das wiederholte Abspielen des Gelernten den Bewusstseinsverlust während des Schlafes notwendig, damit dies frei von Störung durch neu zu enkodierendes Material geschehen kann und der Mensch umgekehrt nicht zu halluzinieren beginnt, wenn ähnliche Prozesse wie beim Enkodieren erneut stattfinden (Diekelmann und Born 2010).

In einer weiteren fMRT-Studie wurde Probanden eine große Zahl Bilder gezeigt und zu verschiedenen Zeitpunkten die korrekte Wiedererkennung geprüft. Die Länge von SWS im Nap direkt nach dem Lernen und vor der ersten Prüfung korrelierte positiv mit dem korrekten und sicheren Erkennen der erlernten Bilder und negativ mit der hippocampalen Beteiligung bei korrektem Erkennen. Im Verlauf von drei Monaten wurde die hippocampale Aktivierung bei korrektem Erkennen immer geringer, dafür stieg die Aktivierung im ventromedialen Präfrontalkortex. Dies ist ein Hinweis auf

Systemkonsolidierung, den Transfer⁶ von Information vom Hippocampus (Kurzzeitspeicher) in den Neokortex (Langzeitspeicher) und die Bedeutung von Schlaf dafür (Takashima et al. 2006). Als alternatives oder ergänzendes Modell wurde die *synaptische Herabregulierung* vorgeschlagen, die von einer Zunahme der Synapsendichte mit vermehrtem Energie- und Platzbedarf im Wachzustand ausgeht, während im nachfolgenden Schlaf die Synapsen gleichmäßig herabreguliert werden. In dem Prozess der Herabregulierung gehen schwache Informationen verloren, die synaptische Homöostase wird wiederhergestellt, das Signal-/Rauschverhältnis wird verbessert und das Gehirn wird wieder aufnahmefähig für neue Enkodierungsprozesse (Tononi und Cirelli 2006; Rasch und Born 2013).

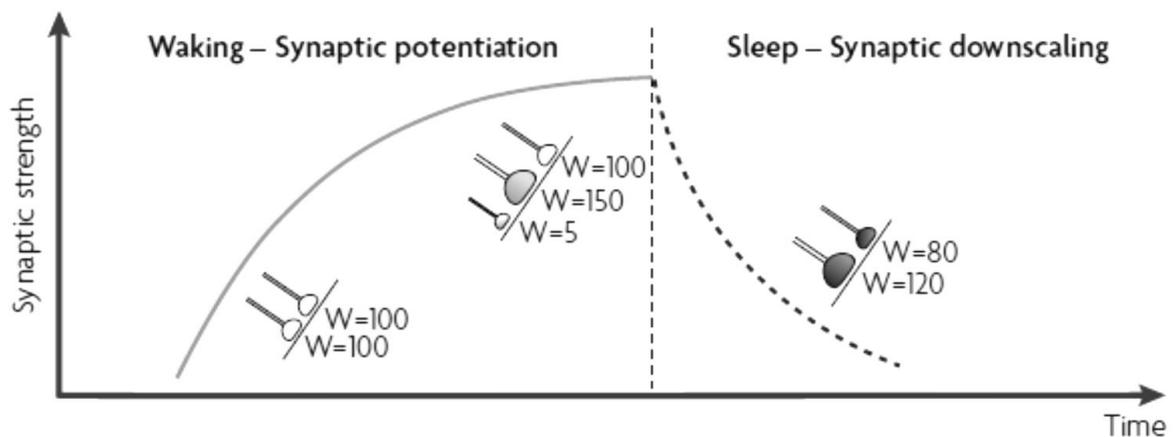


Abb. 4: Konzept der synaptischen Herabregulation, Grafik nach Tononi und Cirelli 2006; modifiziert von Diekelmann und Born 2010

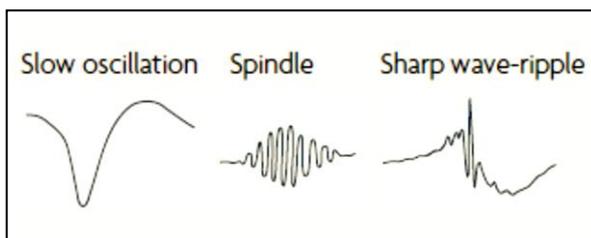


Abb. 5: modifiziert aus Diekelmann und Born 2010

Ein Ansatz diese beiden Konzepte zu vereinen scheint unter Einbeziehung des Leichtschlafes (NREM 2) mit einem Fokus auf die zugrundeliegenden elektrophysiologischen Phänomene möglich: Leichtschlaf macht über 50% des

menschlichen Schlafes aus, und Phänomene, wie langsame Oszillationen (K-Komplexe), Spindeln und Sharp-Wave-Ripples (SWR), die mit Systemkonsolidierung in Verbindung gebracht werden, finden (wenn auch nicht ausschließlich) im Leichtschlaf statt. Darüber

⁶ Der Begriff „Transfer“ wird mittlerweile kritisch gesehen, die tatsächlichen Vorgänge erscheinen deutlich komplexer, vgl. Genzel et al. 2014

hinaus ist Leichtschlaf mit einer größeren hippokampokortikalen sowie kortikokortikalen Interkonnektivität assoziiert und bietet somit bessere Bedingungen für Systemkonsolidierung. Je tiefer der Schlaf wird, desto mehr finden sich Deltawellen als mögliches Korrelat für synaptische Herabregulation. Dieser mehr lokale Vorgang wird auch durch die verringerte Interkonnektivität während SWS befördert. Allerdings scheint es wahrscheinlich, dass wichtige Gedächtnisspuren auch während SWS fortlaufend reaktiviert werden und so der allgemeinen Herabregulation für diese Inhalte entgegengewirkt wird (Genzel et al. 2014).

1.3.3 Schlafspindeln

Schlafspindeln sind EEG-Phänomene des NREM-Schlafes. Sie stellen ein visuelles Merkmal des Stadiums N2 dar, können aber auch in N3 auftreten (Iber et al. 2007). Sie zeichnen sich durch eine Frequenz von etwa 9-15 Hz und eine Dauer von 0,5-3 Sekunden aus und treten häufig in zeitlichem Zusammenhang zu langsamen Oszillationen und K-Komplexen auf (Astori et al. 2013). Teilweise werden langsame und schnelle Spindeln unterschieden: Langsame Spindeln (etwa 9-13 Hz) sind über dem frontalen Kortex lokalisiert, schnelle Spindeln (etwa 13-15 Hz) sind eher verteilt über dem zentralen und parietalen Kortex (Astori et al. 2013). Erschwert wird die Unterscheidung dadurch, dass die Spektren teils überlappen und verschiedene Autoren abweichende Frequenzspektren zugrunde legen. Auch ist noch unklar, ob es sich um zwei grundlegend verschiedene neuronale Prozesse handelt oder lediglich eine Modulation des gleichen Prozesses darstellt (Fogel und Smith 2011; Rasch und Born 2013).

Die Spindeldichte verhält sich invers zur Delta-Aktivität und nimmt im Lauf der Nacht zu, während die Delta-Aktivität (als Marker des homöostatischen Schlafdrucks) kontinuierlich abnimmt (Fogel und Smith 2011). Die Schlafspindeln werden im Thalamus generiert (Krosigk et al. 1993). Dabei gilt der Nucleus Reticularis als „Schrittmacher“ (Steriade et al. 1985).⁷

Die Spindeldichte variiert deutlich nach Geschlecht und von Person zu Person. Intraindividuell hingegen bleibt sie im Vergleich von zwei Nächten ziemlich stabil (Gaillard und Blois 1981). Individuen mit mehr Schlafspindeln können gemäß einer

⁷ Für detailliertere physiologische Grundlagen s. Übersichtsarbeiten: Steriade 2000; Gennaro und Ferrara 2003; Rasch und Born 2013

Studie bei lauter Umgebung besser durchschlafen (Dang-Vu et al. 2010). In verschiedenen Studien wurde ein Zusammenhang von Spindeldichte und allgemeinen kognitiven Fähigkeiten (Bodizs et al. 2005) bzw. von Spindelzahl und „Performance IQ“ (Fogel et al. 2007) festgestellt. Basierend auf diesen und weiteren Studien wurden Spindeln als möglicher „physiologischer Index“ für Intelligenz vorgeschlagen (Fogel und Smith 2011). Neueren Erkenntnissen zufolge besteht allerdings ein deutlicher geschlechtsabhängiger Dimorphismus, so dass lediglich bei Frauen eine moderate positive Korrelation von Intelligenz und einzelnen Spindelparametern feststellbar ist (Ujma et al. 2014).

Vor allem scheinen Schlafspindeln aber auch eine Rolle in der Gedächtniskonsolidierung zu spielen (Übersicht: Fogel und Smith 2011; Rasch und Born 2013): Nach dem Erlernen deklarativer Inhalte kam es bei Probanden einer Studie, verglichen mit der Kontrollbedingung, zu einem signifikanten Anstieg der Spindeldichte, insbesondere innerhalb der ersten 90 Minuten des Schlafes. Darüber hinaus korrelierte die Spindeldichte mit der Abrufleistung vor und nach dem Schlaf. Der Anteil der Schlafstadien hingegen war unverändert (Gais et al. 2002). In einer weiteren Studie wurden Probanden danach eingeteilt, ob es nach einer deklarativen Lernaufgabe (Wortpaare) zu einem Anstieg der Spindelaktivität (enthält die Faktoren Spindeldauer und Spindelamplitude) im Schlaf verglichen mit einer Kontrollnacht kam oder nicht. Nur Probanden mit einem Anstieg der Spindelaktivität zeigten (als Gruppe) am nächsten Morgen eine signifikante Verbesserung der Abrufleistung verglichen mit dem Abend zuvor. Dabei korrelierte die Spindelaktivität in der Nacht nach Lernen zwar signifikant mit der Verbesserung über die Nacht, nicht jedoch mit der absoluten Abrufleistung (Schabus et al. 2004). Auch für Naps konnte eine Zunahme der Spindeldichte nach Lernen gezeigt werden, wobei dieser Effekt allerdings nur eine schwierige deklarative Lernaufgabe (abstrakte Begriffe) betraf und für eine vergleichbar leichtere deklarative Aufgabe (konkrete Begriffe) nicht gezeigt werden konnte. Auch eine positive Korrelation der Veränderung der Abrufleistung mit der Zunahme der Spindeldichte konnte nur bei der schwereren Bedingung gefunden werden (Schmidt 2006). Im Gegensatz zu den bereits zitierten Studien, die Effekte für Spindeln im Stadium 2 NREM fanden (Gais et al. 2002; Schabus et al. 2004), kommt eine jüngere Studie zu dem Ergebnis, dass nur im SWS eine Korrelation zwischen Spindeldichte und Gedächtniskonsolidierung besteht (Cox et al. 2012). Eine umfassende Diskussion der Literatur zu der Rolle der langsamen und schnellen Spindeln für die Gedächtnisfunktion würde an dieser Stelle den Rahmen

sprengen und wird dadurch erschwert, dass die meisten der hier zitierten Studien keine Unterscheidung treffen (Schabus et al. 2004; Gais 2006; Cox et al. 2012). Zusammenfassend kann gesagt werden, dass es deutliche Hinweise für eine positive Korrelation zwischen Spindeln und Konsolidierungsprozessen gibt, wobei diese Korrelation alleine jedoch nicht ausreicht um einen kausalen Zusammenhang nachzuweisen (Rasch und Born 2013; Genzel et al. 2014).

Es gibt auch die Hypothese, dass die Spindeln selbst nicht die aktive Gedächtniskonsolidierung befördern, sondern vielmehr die langsamen Oszillationen und die sehr schnellen Sharp-Wave-Ripples. Allerdings existiert ein Zusammenhang zwischen langsamen Oszillationen und Spindeln, so dass die Spindeldichte als Marker für die Dichte der langsamen Oszillationen fungieren kann (Genzel et al. 2014). Unabhängig von der Funktion in der Systemkonsolidierung, scheinen Spindeln aber lokal eine wichtige Rolle in der synaptischen Plastizität zu spielen (Astori et al. 2013; Genzel et al. 2014). Einen möglichen Ansatz zur Erforschung der Funktion von Spindeln stellt auch deren pharmakologische Manipulationen dar. So konnte in einer Studie unter Verwendung eines kurzwirksamen GABA_A-Agonisten (das Schlafmittel Zolpidem) die Anzahl der Spindeln erhöht werden, was zu einer Verbesserung der Abrufleistung in einem Wortpaar-Assoziationstest führte (Mednick et al. 2013).

1.3.4 Nap und Gedächtnis

Es stellt sich die Frage, ob die positiven Effekte, die für Nachtschlaf und in Nachthälften-Designs gezeigt wurden, auch schon während kurzen Schlafperioden wie Naps nachweisbar sind (Übersicht: Diekelmann et al. 2009; Ficca et al. 2010): In einer Reihe von Studien konnte dieser positive Effekt von Naps auf die deklarativen Gedächtnisinhalte belegt werden (Tucker et al. 2006; Schabus et al. 2005; Lahl et al. 2008). Anschließend stellt sich die Frage, ob die Konsolidierung ab einem bestimmten Punkt gemäß einem Alles-oder-Nichts Prinzip auftritt, oder eher dosisabhängig von der Dauer des Schlafes bzw. einzelner Stadien wie SWS (Diekelmann et al. 2009). Während sich in einer Studie nur ein positiver Effekt fand, wenn auch SWS vorlag (Schabus et al. 2005) und andere zumindest eine positive, wenn auch nicht signifikante Korrelation von SWS zum Lernerfolg zeigen konnten (Tucker et al. 2006; Tucker und Fishbein 2008), kommt eine dritte Gruppe zu der Schlussfolgerung, dass der positive Effekt vermutlich nicht durch ein bestimmtes NREM-Stadium vermittelt wird (Lahl et al. 2008). Schon ein sehr kurzer Nap von 6 Minuten Dauer zeigte einen positiven Effekt verglichen mit der

Wach-Kontrollgruppe, was die Autoren der Studie zu der Vermutung veranlasste, dass schon der Schlafbeginn einen Konsolidierungsprozess anstößt, der dann möglicherweise im Wachzustand weiterläuft. Auch enthalten diese sehr kurzen Naps kein SWS, sondern nur Leichtschlaf. Allerdings zeigte die Studie bei einer Napdauer von 35 Minuten eine signifikante weitere Verbesserung, was die Autoren zu der Vermutung einer nicht kontinuierlich, sondern in diskreten Stufen verlaufenden Progression führte (Lahl et al. 2008). Auch wenn ein direkter Vergleich von Nachtschlaf und Nap bisher nicht durchgeführt wurde, so lässt sich anhand der in Studien angegebenen Daten feststellen, dass schon Naps von 1-2 h Dauer zu einer vergleichbaren Leistungssteigerung in deklarativen Aufgaben führen, wie eine Nachthälfte oder eine vollständige Nacht (Diekelmann et al. 2009). Allerdings muss angemerkt werden, dass zwischen Nachtschlaf und Nap auch Unterschiede bezüglich zirkadianer Faktoren, wie beispielsweise Kortisolsekretion, bestehen, die hier modulierend wirken können (Tucker et al. 2006). Auch bleibt die Frage offen, inwieweit der Effekt von bleibender Dauer ist, oder ob er mit dem Schlaf der darauffolgenden Nacht auf das gleiche Niveau gebracht wird wie ohne Nap.

1.4 Testing Effekt

A curious peculiarity of our memory is that things are impressed better by active than by passive repetition. I mean that in learning by heart (for example), when we almost know the piece, it pays better to wait and recollect by an effort from within, than to look at the book again. If we recover the words in the former way, we shall probably know them the next time; if in the latter way, we shall very likely need the book once more.

William James (1890)

Bei Universitätsstudenten ist das wiederholte Lesen von Texten die am häufigsten angewandte Lernstrategie und deutlich populärer als Formen der Selbsttestung wie freies Erinnern oder Karteikarten. Diese gelebte Praxis steht in klarem Widerspruch zu wissenschaftlichen Erkenntnissen, dass wiederholtes Lesen kaum Vorteile bringt, während Testing (wiederholter Abruf des Gelernten aus dem Gedächtnis) eine deutliche Verbesserung der langfristigen Behaltensleistung bewirkt (Karpicke et al. 2009; Roediger und Butler 2011). Dieser positive Effekt wurde schon zu Beginn des letzten Jahrhunderts beschrieben (u.a. Abott 1909). Trotzdem blieb der Testing Effekt lange eher unbekannt und Glover fasste die Situation 1989 treffend mit dem Titel “The ‘Testing’ Phenomenon: Not Gone but Nearly Forgotten” zusammen (Glover 1989). Im neuen Jahrtausend hat die Forschung jedoch deutlich an Fahrt aufgenommen, so dass in den vergangenen Jahren einige neue Erkenntnisse zum Testing Effekt und seinen möglichen Anwendungsgebieten gewonnen wurden (Roediger und Butler 2011).

1.4.1 Stand der Forschung

Im Folgenden werden die Ergebnisse einer Studie von Karpicke und Roediger (2008) zum Testing Effekt vorgestellt: Sie untersuchten an Studenten verschiedene Methoden des Vokabellernens und ihren Effekt auf die langfristige Behaltensleistung. Die verwendeten Wortpaare (Suaheli – Englisch, z.B. adhama – honor) wurden wiederholt gezeigt und geprüft, bis sie erfolgreich beherrscht wurden. Sobald ein Wortpaar einmal korrekt gewusst wurde, wurde es einer von vier Übungsbedingungen zugeordnet: Es wurde (1) weiterhin gezeigt und geprüft, (2) nicht mehr gezeigt aber weiterhin geprüft, (3) weiterhin gezeigt aber nicht mehr geprüft, oder (4) weder gezeigt noch geprüft. Während die Wortpaare in allen vier Bedingungen etwa gleich gut erlernt wurden (nach vier Durchgängen wurden alle Wortpaare beherrscht), gab es bei einer erneuten Prüfung

nach einer Woche sehr deutliche Unterschiede in der Behaltensleistung zwischen den Bedingungen mit wiederholtem Abruf (1, 2) und den Bedingungen ohne wiederholten Abruf (3, 4).

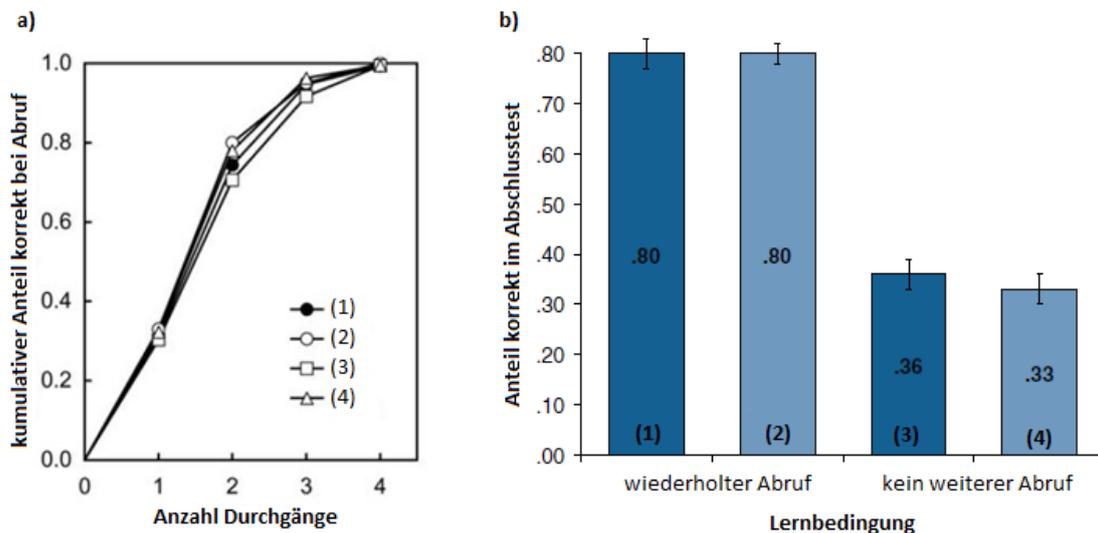


Abb. 6: a) Erlernen in etwa gleich in allen vier Bedingungen, Grafik modifiziert nach (Karpicke und Roediger 2008); b) Differenz von 4 Standardabweichungen zwischen Bedingungen mit und solchen ohne wiederholten Abruf (=Testing), Grafik modifiziert nach (Roediger und Butler 2011)

In einer anderen Studie (ebenfalls mit Wortpaaren) konnte gezeigt werden, dass längere Intervalle zwischen den einzelnen Übungstests (6 min vs. 1 min) eine deutlich höhere Behaltensleistung eine Woche später bedingen. Je häufiger der Abruf wiederholt wird, desto besser die langfristige Behaltensleistung. Das gilt allerdings nur für Abruf mit ausreichender Schwierigkeit (in diesem Fall 6 Minuten Intervall). Bei dem einminütigen Intervall hingegen führten bis zu zehn Wiederholungen langfristig zu keiner sichtbaren Verbesserung. Diese Ergebnisse bestätigen die Hypothese, dass eine gewisse mentale Anstrengung beim Abruf sich positiv auf das langfristige Behalten auswirkt (Pyc und Rawson 2009). Auch nach einem Monat zeigt ein Test direkt nach Lernen noch positive Effekte (Butler und Roediger 2007). Während Testing auch ohne Feedback schon zu einer deutlichen Steigerung in der Behaltensleistung führt, bringt Feedback häufig einen zusätzlichen Gewinn, insbesondere indem falsch gemerkte Elemente korrigiert werden können. Besonders bei Aufgaben, die mit Falschantworten als Ködern arbeiten (z.B. Multiple-Choice Fragen) ist Feedback wichtig um das Erlernen falscher Fakten zu verhindern. Es gibt Hinweise, dass zeitlich verzögertes Feedback einen noch größeren Nutzen bringt, als Feedback direkt nach der Beantwortung der Frage (Butler und Roediger 2008; Roediger und Butler 2011).

Die positiven Effekte von wiederholtem Abruf konnten für eine Vielzahl verschiedener Gedächtnisinhalte gezeigt werden (Übersicht: Dunlosky et al. 2013): für naturwissenschaftliche Texte, wo es um das Verstehen und Behalten von Konzepten geht (Karpicke und Blunt 2011), für Wortpaare, Vokabellisten, Definitionen, aber auch für das korrekte Benennen von Objekten und Personen und deren Lokalisation im Raum (Dunlosky et al. 2013). Darüber hinaus konnte auch ein Testing Effekt für nicht-deklarative Aufgaben wie den Erwerb von praktischen Fähigkeiten gezeigt werden. Medizinstudenten lernten in einem Kurs das Notfallmanagement bei Patienten mit Herzstillstand. Wurde zu Ende des Kurses das Gelernte getestet war die Leistung zwei Wochen später besser als in einer Kontrollgruppe (Kromann et al. 2009). Der Testing Effekt funktioniert sowohl mit offenen Fragen, als auch mit Multiple-Choice-Fragen, Lückentexten und sogar wenn es ermöglicht wird im Übungstest die Lösung nachzulesen. Dabei können Übungstests und Abschlusstest auch ein unterschiedliches Format verwenden, der Effekt bleibt trotzdem bestehen. Der Nutzen beschränkt sich hierbei nicht allein auf die exakten abgerufenen Inhalte, sondern es zeigt sich auch eine Verbesserung in Transferfragen, d.h. in der Anwendung der erlernten Konzepte auf neue Fragestellungen (Dunlosky et al. 2013). Testing führt sogar zu einer gesteigerten Behaltensleistung bei solchen Inhalten, die nicht selbst getestet wurden, aber eine inhaltliche Verwandtschaft zu getesteten Elementen aufweisen (Chan et al. 2006). Das Einstreuen von Zwischentests während längerer Lernsitzungen schützt die später gelernten Inhalte vor proaktiver Interferenz durch die kurz davor gelernten Elemente (Szpunar et al. 2008). Darüber hinaus führt wiederholter Abruf zu einer besseren Organisation des Gelernten (Zaromb und Roediger 2010).

Entsprechend dieser Erkenntnisse kann der Testing Effekt von großem Nutzen für Schüler und Studenten sein (Roediger et al. 2011; Larsen et al. 2008; McDaniel et al. 2007), aber auch Patienten, die wegen Multipler Sklerose an einer eingeschränkten Gedächtnisleistung leiden, können von Testing profitieren (Sumowski et al. 2013). Allerdings kann Testing auch negative Effekte haben, z.B. werden nicht gelernte aber inhaltlich verwandte Wörter häufiger fälschlicherweise angegeben (McDermott 2006) und die Beeinflussbarkeit von Augenzeugenberichten wird mit dem wiederholten Wiedergeben des Erlebten verstärkt (Chan und LaPaglia 2011).

1.4.2 Erklärungsansätze

Die Vermutung, dass der Testing Effekt lediglich auf einer größeren Exposition mit dem Gelernten beruht, konnte in einer Reihe Studien widerlegt werden, die den Abrufversuchen eine vergleichbare Anzahl von Reexpositionen mit dem Lernstoff gegenüber stellten (Roediger und Karpicke 2006a; Karpicke und Roediger 2008; Übersicht: Roediger und Karpicke 2006b). Auch liefert dieser Ansatz keine Erklärung dafür, dass in Studien direkt nach der Lernphase die Gruppen mit wiederholter Exposition besser abschnitten, nach einem zeitlichen Intervall hingegen die Gruppen mit wiederholtem Abruf bessere Ergebnisse zeigten (Roediger und Karpicke 2006a, 2006b).

Es werden in der Literatur mehrere Ansätze diskutiert, die sich dabei nicht gegenseitig ausschließen müssen, sondern vielmehr in Ergänzung zum Effekt beitragen können (Übersicht: Roediger und Karpicke 2006b; Roediger und Butler 2011): Ein Erklärungsansatz besagt, dass der Abruf als solcher zu einer Verstärkung der Gedächtnisspur oder zu einer größeren Anzahl von Abrufpfaden führt. Dabei spielt die Abrufanstrengung eine wichtige Rolle. Ähnlich der Verarbeitungstiefe bei der Enkodierung, führt eine „erstrebenswerte Schwierigkeit“ des Abrufs, z.B. durch längere Intervalle zwischen den Tests, zu tieferegreifenden Prozessen und damit zu größerem Nutzen (Roediger und Karpicke 2006b; Pyc und Rawson 2009). Einen theoretischen Erklärungsansatz liefert hierbei das Konzept von Bjork und Bjork (1992): Sie unterscheiden die Stärke der Speicherung (im Langzeitgedächtnis) und die Stärke der Abrufbarkeit (aktuelle Zugänglichkeit der Erinnerung). Elemente können noch gespeichert, aber nicht mehr abrufbar sein (bspw. eigene alte Telefonnummer). Elemente die schwer abrufbar sind, profitieren nach diesem Modell sehr stark für die Zukunft von einem erfolgreichen Abruf, während gut beherrschte und leicht abrufbare Elemente durch den erfolgreichen Abruf kaum profitieren. Ein weiterer Ansatz argumentiert basierend auf dem Konzept des „Transfer Appropriate Processing“ (Morris et al. 1977), dass Testing bessere Ergebnisse liefert, da das Format der Lernphase (ein Wort wird gezeigt und das zweite Wort des Wortpaares muss aus dem Gedächtnis abgerufen werden) identisch mit dem eigentlichen Test am Ende ist und daher eine bessere Vorbereitung auf diese Abfrageform darstellt als unspezifisches Wiederholen (Roediger und Butler 2011). Ein weiterer Aspekt ist die Effektivität von Mediatoren: Häufig werden beim Lernen Mediatoren (z.B. Schlüsselworte oder Eselsbrücken) entwickelt und verwendet. Findet zwischen den Lernphasen, in denen die Wortpaare präsentiert werden, wiederholt Testing statt, so wird der mit dem Wortpaar verbundene Mediator später besser erinnert und führt häufiger zu

Erinnerung an die richtige Lösung. Auch kommt es in der bei Testing häufiger dazu, dass Mediatoren verändert und durch effektivere ersetzt werden. (Pyc und Rawson 2010). Darüber hinaus ist noch das Bifurkationsmodell zu nennen⁸: Dieses Modell geht im Gegensatz zu anderen Modellen nicht von einer Verlangsamung der Vergessensrate nach Testing, sondern von einer selektiven Stärkung der erfolgreich abgerufenen Elemente aus. Die einzelnen Gedächtnisinhalte sind von vornherein verschieden stark gespeichert. Das Wiederholen der Inhalte führt zu einer gleichmäßigen und schwachen Verstärkung *aller* Elemente, während bei wiederholtem Abruf ohne Feedback erfolgreich abgerufene Inhalte deutlich verstärkt werden und die nicht erfolgreich abgerufenen Inhalte unverändert bleiben. So kommt es nach Testing zu einer Zweiteilung der Gedächtnisinhalte, zu der sog. Bifurkation. Alle Elemente werden gleich schnell vergessen unabhängig von der verwendeten Lernmethode. Der Abschlusstest am Ende misst nicht die absolute Gedächtnisstärke aller Elemente, sondern die Anzahl der Elemente, die sich oberhalb der Schwelle befinden, wo sie einem Abrufversuch des Bewusstseins gerade noch zugänglich sind. Da erfolgreich abgerufene Elemente besonders verstärkt wurden dauert es verglichen mit wiederholten Elementen, die nur schwach verstärkt wurden, länger bis sie im Rahmen des kontinuierlichen Vergessensprozesses unter die Abrufbarkeitsschwelle fallen. Diese Zusammenhänge sind in Abb. 7 noch einmal verdeutlicht.

⁸ Nachfolgende Ausführungen zum Bifurkationsmodell gemäß Kornell et al. 2011

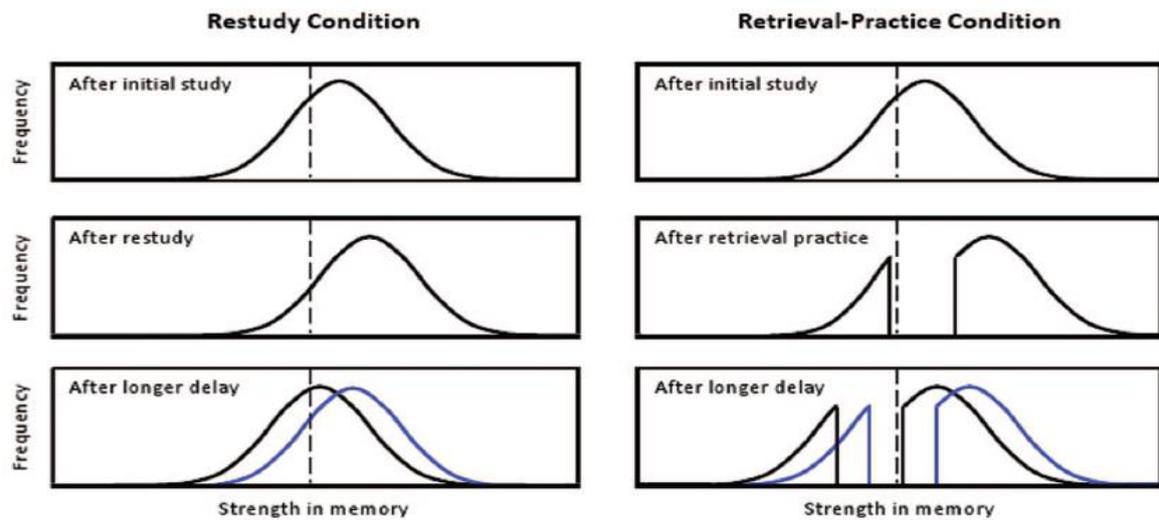


Abb. 7: Beispielhafte Darstellung der Gedächtnisstärke von einer großen Anzahl erlernter Elemente. Normalverteilung angenommen. Die gestrichelte Linie symbolisiert die Abrufbarkeitsschwelle, Elemente rechts davon können erinnert werden, Elemente links sind nicht abrufbar. Die linke Spalte zeigt wiederholtes Einprägen, die rechte wiederholten Abruf ohne Feedback, jeweils zu mehreren Zeitpunkten (von oben nach unten). Es ist deutlich erkennbar, dass sich links nach kurzer Zeit mehr Elemente im zugänglichen Bereich befinden, nach einem längeren Intervall jedoch sind rechts (Testbedingung) mehr Elemente oberhalb der Abrufbarkeitsschwelle. Grafik aus Bäuml et al. 2014, modifiziert nach Kornell et al. 2011

2 Zielsetzung der Arbeit

Ein spannender Aspekt des Testing Effekts ist seine Zeitabhängigkeit: Während direkt im Anschluss an die Enkodierphase kein deutlicher Unterschied zwischen den Bedingungen sichtbar ist oder sogar die Bedingung mit wiederholter Präsentation des Lernstoffs bessere Ergebnisse liefert, führt wiederholter Abruf langfristig zu deutlich besseren Ergebnissen (u.a. Roediger und Karpicke 2006a; Karpicke und Roediger 2008; Toppino und Cohen 2009). Einen möglichen Erklärungsansatz bietet hier die Bifurkationshypothese (s. oben). In der vorliegenden Arbeit soll aber ein anderer Zusammenhang beleuchtet werden: Schlaf spielt eine wichtige Rolle bei Gedächtnisprozessen und neben der erfolgreichen Enkodierung ist die aktive Gedächtniskonsolidierung (die besonders im Schlaf stattfindet) von zentraler Bedeutung für die Permanenz und Abrufbarkeit von Gedächtnisinhalten. Dabei kann das Gehirn während des Schlafes selektiv Inhalte von künftiger Relevanz verstärkt konsolidieren (u.a. Wilhelm et al. 2011). Das führt zu der Frage, ob das verzögerte Auftreten der positiven Auswirkung von wiederholtem Abruf in der Notwendigkeit der schlafabhängigen Konsolidierung für das Auftreten dieses Effekts begründet liegt. Ausgangspunkt der eigenen Untersuchung ist die Annahme, dass nicht allein das Verstreichen von Zeit zu diesem Ergebnis führt, sondern dass der Schlaf, der während dieses Intervalls stattfindet, eine entscheidende Rolle spielt. Es ist vorstellbar, dass Wissen, welches durch wiederholten Abruf eingeübt wurde, im Schlaf anders verarbeitet und konsolidiert wird als Wissen, das durch wiederholte Präsentation gelernt wurde.

Um diesen Zusammenhang von Testing Effekt und schlafabhängiger Gedächtniskonsolidierung zu beleuchten, wurde eine randomisierte Studie an gesunden männlichen Probanden durchgeführt. Im Rahmen dieser Studie sollen folgende Fragestellungen beantwortet werden:

- Können die vorbeschriebenen positiven Effekte von Schlaf auf Gedächtnis und von wiederholtem Abruf auf das langfristige Behalten an einem großen Probandenkollektiv bestätigt werden?
- Gibt es eine Interaktion zwischen den Effekten von Schlaf und von wiederholtem Abruf auf die Gedächtnisleistung? Beeinflusst Schlaf die Ausprägung des Testing Effekts?

- Gibt es Korrelationen zwischen Schlafdauer, einzelnen Schlafstadien oder Spindeldichte mit dem individuellen Lernerfolg? Gibt es einen Unterschied in diesen Schlafparametern oder dem Schlaf-EEG zwischen dem Nap nach Lernen und dem Kontrollnap?

3 Material und Methoden

3.1 Studienprobanden

An der Studie nahmen insgesamt 80 gesunde männliche Studenten im Alter zwischen 18 und 30 Jahren teil. Zahlreiche Fachrichtungen waren vertreten, am häufigsten waren Medizinstudenten (16) und Physikstudenten (12). Alle Probanden hatten Deutsch als Muttersprache, um an diesem Punkt gleiche Ausgangsbedingungen bei der Lernaufgabe zu gewährleisten. Als Ausschlusskriterien galten bei der Probandenrekrutierung Schlafstörungen, psychische Erkrankungen, Schichtarbeit, transmeridiane Flüge im vergangenen Monat und ein extremer Abendtyp mit regelhaft spätem Zubettgehen nach ein Uhr nachts. Auch sollten die Probanden keinen täglichen Mittagsschlaf halten, keine Drogen oder starken Medikamente einnehmen, sowie Kaffee (max. 2 Tassen/Tag), Alkohol (max. zweimal/Woche, nicht am Vorabend der Testung) und Zigaretten (max. 5 Stück/Tag) nur in Maßen genießen. Die Rekrutierung der Probanden erfolgte im Freundes- und Bekanntenkreis, mittels Aushängen an der Universität und über den offiziellen Uni-Verteiler. Für ihre Teilnahme erhielten die Probanden eine Aufwandsentschädigung von insgesamt 120 Euro für drei Termine.

Bevor neue Probanden in die Studie eingeschlossen werden konnten, mussten sie per E-mail drei Fragebögen beantworten: Einen Fragebogen zum Chronotyp (D-MEQ; Horne und Ostberg 1976), einen Fragebogen zum gegenwärtigen Lebensgefühl und Ausschluss Depression (BDI-V; Schmitt und Maes 2000; Schmitt et al. 2006) und einen Fragebogen zur Schlafqualität (PSQI; Buysse et al. 1989). Mithilfe dieser Fragebögen konnten Probanden mit einem stark abweichenden Schlafrhythmus, einer Schlafstörung oder einer Depression, die sich in der Regel auf die Schlafqualität auswirkt (Dresler et al. 2014), identifiziert und von der Teilnahme ausgeschlossen werden.

Vor ihrem ersten Termin im Labor erhielten alle Teilnehmer der Studie eine ausführliche mündliche und schriftliche Probandenaufklärung. Diese enthält u.a. eine Beschreibung des Untersuchungsablaufs und Angaben zum Datenschutz. Ein zustimmendes Votum der zuständigen Ethikkommission der Medizinischen Fakultät der Universität München zur Studie liegt vor.

3.2 Studiendesign

3.2.1 Weitere Fragestellungen

Neben den Versuchen zur Interaktion von Schlaf und Testing bearbeiteten die Studienprobanden auch Aufgaben im Rahmen weiterer Fragestellungen, die nicht Thema dieser Arbeit sind⁹ und an dieser Stelle nur der Vollständigkeit halber Erwähnung finden sollen: Zum einen ging es hier um die Frage, wie sich Schlaf auf Kreativität bei Produktentwicklungsaufgaben auswirkt (Marisch 2015) und ob Pseudogruppen aus drei Einzelpersonen in Summe kreativer (gemessen an Qualität und Quantität von Produktentwicklungen zu einem vorgegebenen Thema) sind als Dreiergruppen, die miteinander kooperieren. Zum anderen wurde die Auswirkung von Schlaf auf das gefühlte Eigentum und Risikoverhalten im Umgang mit Geld untersucht. Erhält ein Mensch unerwartet Geld, neigt er dazu dieses Geld schneller und leichter auszugeben¹⁰, es hat in seinem Empfinden einen anderen Wert als sein eigenes Geld. Nach einer Weile ist es jedoch in sein gefühltes Eigentum übergegangen und wird ebenso behandelt. Mithilfe einer Wettaufgabe, bei der Probanden zwischen risikoreichen und risikoarmen Varianten wählen konnten wurde versucht zu ermitteln, ob Schlaf einen Einfluss darauf hat, wie schnell unerwartet erhaltenes Geld Teil des gefühlten Eigentums wird. Auch wurde Probanden die Möglichkeit gegeben vom Gewinn einen Anteil ihrer Wahl für einen wohlthätigen Zweck zu spenden. Beide zusätzliche Fragestellungen wurden vor der Enkodierungsphase bzw. nach der Abrufphase durchgeführt, der für die vorliegende Arbeit wesentliche Versuchsablauf Enkodierung-Nap-Abruf war somit von diesen zusätzlichen Aufgaben unangetastet.

3.2.2 Ablauf der Studientage

Jeder Proband kam für drei verschiedene Versuchsnachmittage in das Schlaflabor des Instituts. Die Reihenfolge der einzelnen Studientage war randomisiert und zwischen den einzelnen Versuchstagen lag ein mindestens zweiwöchiges Intervall. In der Woche vor dem Versuchstag mussten die Probanden ein Schlafprotokoll führen. Am Vorabend der

⁹ Zum aktuellen Zeitpunkt ist nur eine Dissertation abgeschlossen (Marisch 2015), zwei weitere sind in Arbeit.

¹⁰ Für einen kleinen Einblick in die Thematik der Windfall Gains sei folgender Artikel empfohlen: Arkes et al. 1994

Testung sollten sie spätestens um 1 Uhr nachts schlafen gehen und am Versuchstag selbst mussten sie um 7 Uhr morgens aufstehen. Durch die dementsprechend lange Wachphase vor den Versuchen und die Vermeidung möglicherweise überlanger REM-reicher Schlafstunden am Morgen sollte ein hinreichender Schlafdruck gewährleistet werden, um im Nap auch REM und Tiefschlaf erreichen zu können. Die Versuche begannen pünktlich um 13 Uhr und dauerten je nach Bedingung und Bearbeitungsgeschwindigkeit zwischen drei und fünf Stunden. Es gab für die beiden parallelen Versuchstage je parallel gestaltete Aufgaben, deren Zuordnung zu Wach oder Nap ebenso wie die Reihenfolge der Versuchstage unabhängig voneinander durchpermutiert wurden.

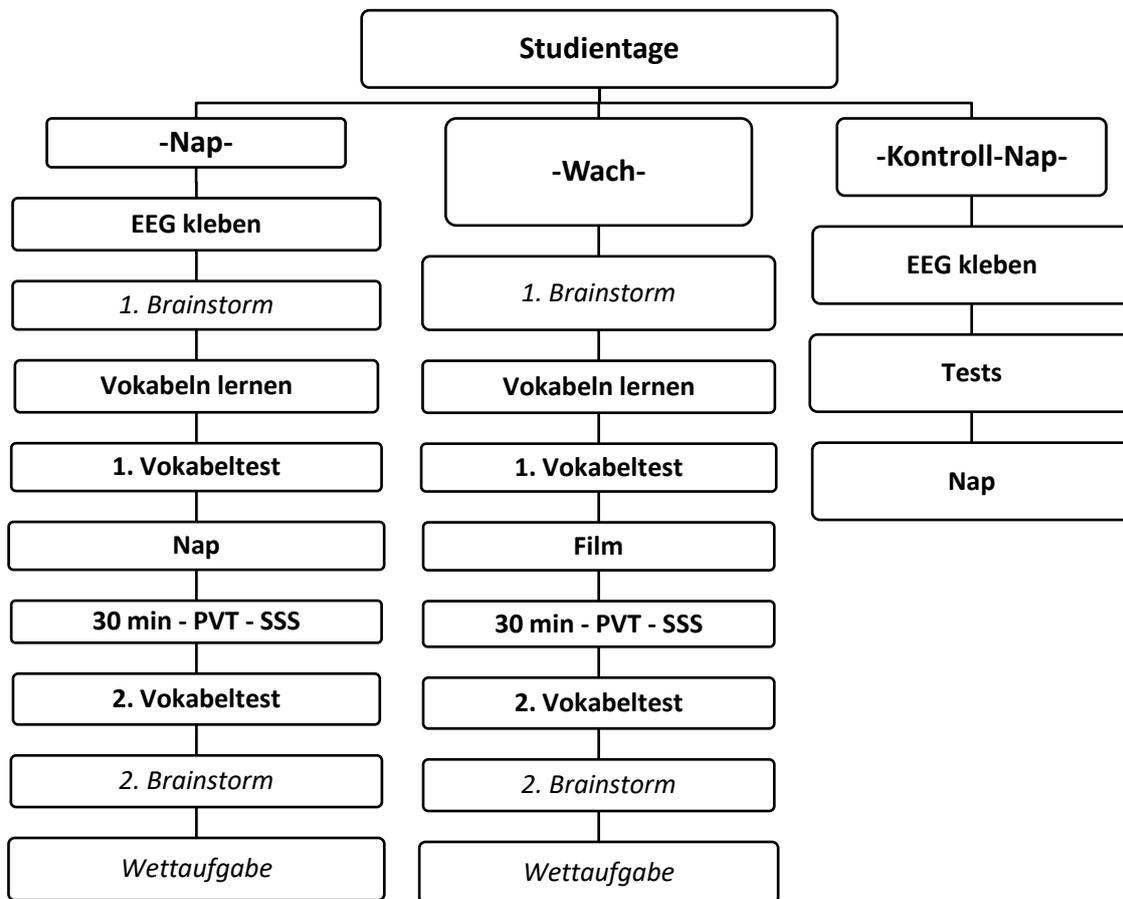


Abb. 8: Ablauf der einzelnen Studientage; Elemente mit Relevanz für die vorliegende Arbeit sind durch fette Schrift hervorgehoben.

3.2.2.1 Ablauf der Nap- und Wachbedingung

In der Nap-Bedingung wurden zu Beginn des Versuchstages die Elektroden für die polysomnographische Ableitung angelegt (näheres dazu s. 3.4). In der Wachbedingung

entfiel dieser Teil. Dadurch kam es insgesamt zu einem leichten zeitlichen Unterschied zwischen den ansonsten parallel gehaltenen Tagen (s. dazu auch 5.4) Danach liefen beide Bedingungen, mit Ausnahme des Nap, identisch ab. Nach einer kurzen Einführung mit Übersicht über den Ablauf des jeweiligen Studientages begannen die Probanden mit der Brainstorming-Aufgabe. Hier sollten zu zwei vorgegebenen Produktkategorien möglichst viele und neuartige Produkte entwickelt werden. Die Hälfte der Probanden bearbeitete diese Aufgabe in Dreiergruppen, während die andere Hälfte es in Einzelarbeit durchführte. Zum Schluss wurden die Erfindungen von den Probanden durchnummeriert, je nachdem welche sie für die besten und kreativsten Produkte hielten. Als nächstes folgte die Lernaufgabe, in der den Probanden zunächst alle Wortpaare einmal präsentiert wurden und anschließend ein Teil der Probanden durch wiederholte Präsentation die Wortpaare lernte, während der andere Teil der Probanden wiederholt abgeprüft wurde und nur bei falscher Antwort das Wortpaar erneut gezeigt bekam. Im Anschluss an die Lern- oder Übungsphase wurde der Lernerfolg in einem Vokabeltest abgefragt (für eine detaillierte Beschreibung der Lernaufgabe s. 3.3). Nun folgte der abweichende Teil zwischen beiden Versuchstagen: In der Nap-Bedingung gingen die Probanden in das Schlaflabor und nachdem die Polysomnographie betriebsbereit und eine Bioeichung durchgeführt war durften sie 90 Minuten schlafen. Dann wurden sie wieder geweckt, die EEG Elektroden wurden entfernt und sie konnten sich kurz im Bad frisch machen, bevor die Versuche fortgeführt wurden. Den Probanden der Wachbedingung wurde ein Naturfilm („Unsere Erde“) von 95 Minuten Länge gezeigt. Um eine zeitliche Parallelität zu der Nap-Bedingung in Bezug auf das Intervall zwischen den Versuchsblocks vor und nach dem Film/Nap zu gewährleisten, wurde eine entsprechende Pause eingehalten, bevor die Versuche fortgeführt wurden. Als erste Aufgabe des zweiten Versuchsblocks (ab hier waren beide Versuchsbedingungen wieder identisch) mussten alle Probanden einen Psychomotor Vigilance Task (Dinges und Powell 1985; Loh et al. 2004) am Computer durchführen, mit dem Aufmerksamkeitslücken und Reaktionsgeschwindigkeit auf visuelle Reize über fünf Minuten gemessen wurden. Auf einem Fragebogen (Stanford Sleepiness Scale; Hoddes et al. 1973) kreuzten die Probanden an, wie müde sie sich aktuell fühlten (von 1 – *ganz wach* bis 7 – *am Einschlafen*). Es folgte ein weiterer Vokabeltest. Danach wurden noch einmal Produkte für die gleichen beiden Kategorien wie vorher entwickelt, diesmal allerdings mit weniger Zeit. Dann wurden die Produkte (sowohl die vor dem Nap / Film als auch die ganz neu entwickelten) wieder durchnummeriert, je nachdem welche die Probanden für die besten und innovativsten hielten. Als letzter großer Versuchsblock folgte die Wettaufgabe. Dazu erhielten die

Probanden zehn Euro (manche zu Beginn des Versuchstages, der Rest direkt vor Beginn der Aufgabe), mit denen sie Wetten abschließen mussten. Hierzu mussten sie sich zwischen zwei verschiedenen risikoreichen Varianten entscheiden. Auch mussten sie vorher festlegen, ob und wie viel ihres potentiellen Gewinns sie spenden würden. Dann wurde mithilfe zweier Würfel eine Wette ausgewählt und die Option, die der Proband für diese Wette gewählt hatte, wurde ausgespielt. Der Proband bekam den vollen Endbetrag ausgezahlt und gab den angegebenen Spendenbetrag dafür in eine Spendenbox. Vier Tage nach der Testung mussten die Probanden online erneut den Vokabeltest durchführen.

3.2.2.2 Ablauf der Kontrollnap-Bedingung

In der Kontrollnap-Bedingung ging es darum, im Intra-Subjekt-Vergleich zu prüfen, wie der Schlaf durch die Kreativitäts- und Lernaufgaben verändert ist. Dazu kamen die Probanden einen weiteren Tag ins Labor, um dort zu schlafen, ohne dass Versuche mit potentiell schlaf-abhängigen Aufgaben durchgeführt wurden. Allerdings wurden an diesem Tag einige Fragebögen und Tests bearbeitet: Ein Fragebogen zur Einschätzung der eigenen Kreativität (basierend auf CSQ-R; Kumar et al. 1997) und zur Händigkeit (EHI; Oldfield 1971), sowie folgende Tests zur Messung der Intelligenz:

- MWT: Hier wurde eine verkürzte Form des MWT (Mehrfach-Wortschatz-Intelligenz-Test) eingesetzt, der nur noch die schwierigeren Auswahlpunkte (15-37) enthielt. Bei diesem Test muss der Proband von fünf angebotenen Worten, dasjenige auswählen welches in der deutschen Sprache vorkommt. Der MWT misst das allgemeine Intelligenzniveau (Lehrl 2005). Er dient im Rahmen der Studie auch zur Kontrolle einer ausreichenden Beherrschung der deutschen Sprache, so dass dadurch keine Einschränkung in der Lernaufgabe zu erwarten ist.
- CFT: Bei diesem Test handelt es sich um einen klassischen Test zur Messung der fluiden Intelligenz. Er wird altersabhängig ausgewertet und liefert als Ergebnis einen IQ-Wert (Weiss und Weiss 2006).

3.3 Lernaufgabe

3.3.1 Entwicklung, Validierung und Programmierung der Lernaufgabe

Um die Interaktion von Schlaf und Testing Effekt auf die langfristige Behaltensleistung zu untersuchen, wurde eine Reihe von 36 Wortpaaren (Deutsch – Fantasiensprache)

erstellt, welche die Probanden erlernen mussten. Um auch subtile Effekte auf die Behaltensleistung erfassen zu können, mussten Decken- und Bodeneffekte bei der Lernaufgabe vermieden werden und es wurde eine Erfolgsquote von etwa 70% der zu erlernenden Wortpaare zu Ende der Lernphase angestrebt. Da keine der in vergleichbaren Studien verwendeten Wortlisten diese Bedingung erfüllte, wurde eine eigene Liste von Wortpaaren bestehend aus einem deutschen Wort und einem Fantasiewort, das den Falschantworten des MWT entnommen war, entwickelt. In einer Vorstudie wurden die Wortpaare validiert und daraus zwei Listen zu jeweils 36 Wortpaaren erstellt, wobei die Listen im Mittelwert und der Varianz der Erfolgsquote so identisch wie möglich gemacht wurden. In einer zweiten Vorstudie wurden die Lernbedingung und die Abrufbedingung vergleichend getestet und die Anzahl der Wiederholungen sowie die Präsentationszeiten so angepasst, dass beide Bedingungen zu Ende der Lernphase einen möglichst identischen Lernerfolg bewirken.

Für die eigentliche Studie wurden die einzelnen Elemente der Lernaufgabe mithilfe von E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA, USA) programmiert und die Ergebnisse von der Software aufgezeichnet.

3.3.2 Ablauf der Lernaufgabe

Jeweils 40 der insgesamt 80 Probanden absolvierten entweder die Lern- oder die Testbedingung. Vor Beginn der Lernaufgabe erhielten die Probanden eine mündliche Einführung vom Versuchsleiter mit dem Hinweis, dass die Verwendung von Mediatoren als Gedächtnisstütze beim Lernen hilfreich sein kann¹¹. Nach dem Start des Programmes und Angabe der Probandeninfo erschien ein kurzer Einführungstext und danach wurden die Wortpaare für jeweils 5 Sekunden in schwarzer Schrift auf weißem Hintergrund angezeigt. Nachdem alle 36 Wortpaare einmal präsentiert worden waren, teilte sich das weitere Vorgehen je nach Bedingung. Den Probanden der Lernbedingung wurden die gleichen 36 Wortpaare in vier weiteren Präsentationszyklen in randomisierter Reihenfolge gezeigt. Den Probanden der Testbedingung hingegen wurde nur das deutsche Wort des Wortpaares gezeigt und sie mussten per Tastatureingabe das dazugehörige Wort der Fantasiessprache tippen. War die Angabe korrekt, wurde direkt mit der Abfrage des

¹¹ Dieser Hinweis erfolgte, um eine deutliche Verbesserung der Leistung vom ersten zum zweiten Termin zu vermeiden, welche erfolgen könnte, wenn derartige Strategien erst bei dem zweiten Termin angewendet würden

nächsten Wortpaares fortfahren. Bei inkorrekt eingabe wurde das gesamte Wortpaar noch einmal angezeigt¹². In der Testbedingung wurden drei Zyklen durchlaufen, in denen jeweils alle 36 Wortpaare in randomisierter Reihenfolge abgefragt wurden. Nach Abschluss dieser Lern- und Übungsphase erfolgte in beiden Bedingungen ein Test, in dem alle 36 Wortpaare abgefragt wurden. Auch hier wurde das deutsche Wort angezeigt und das entsprechende Fantasiewort musste eingegeben werden, allerdings erfolgte kein Feedback, ob die Eingabe korrekt war und keine erneute Präsentation des gesamten Wortpaares. Dieser erste Test (Baseline) stellt den Ausgangswert für die weiteren Tests und die darauf basierenden Berechnungen dar. Der zweite Test erfolgte direkt im Anschluss an den Nap bzw. den Film. Der dritte und finale Test (Retest) fand vier Tage nach dem Versuchstag statt. Hierfür erhielten die Probanden eine E-Mail mit Link auf eine Internetseite, wo sie diesen Test online bearbeiten konnten. Dies musste am vierten Tag nach der Testung zwischen 11 und 15 Uhr erfolgen.

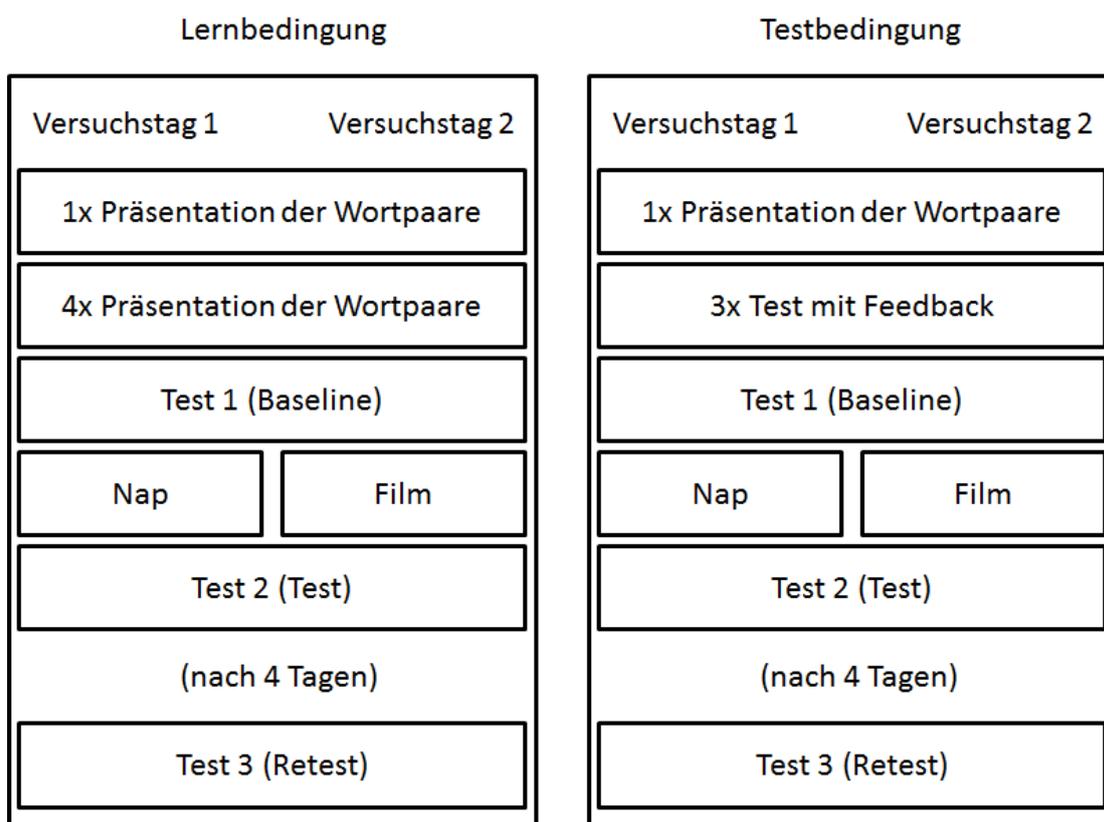


Abb. 9: Ablauf der Lernaufgabe – links für einen Probanden der Lernbedingung, rechts für einen Probanden der Testbedingung

¹² Aufgrund eines Programmierfehlers wurden hier manchmal 3 Sekunden und manchmal 5 Sekunden Feedback gegeben. Statistisch gesehen blieb dieser Fehler ohne größere Folgen, für weitere Details s. Diskussion

3.4 Polysomnographie

3.4.1 Schlafableitung

Die Polysomnographie wurde mittels eines digitalen 12-Kanal-Schreibers aufgenommen, abgespeichert und auch analysiert (Comlab 32 Digital Sleep Lab, Brainlab V 3.3 Software, Schwarzer GmbH, München). Die Anbringung der Kopfelektroden für das EEG erfolgte gemäß dem standardisierten internationalen 10-20-System. Von den in Abb. 10 gezeigten Kopfelektroden wurden N, A1, A2, C3, C4, F3, F4, O1, O2 geklebt. N ist die Erdung, die restlichen Kopfelektroden einer Seite werden im Rahmen einer bipolaren Ableitung gegen A (Mastoid) der Gegenseite geschaltet, gefiltert für einen Bereich von 0,5 bis 70 Hz. Die Abtastrate beträgt 250 Hz. Im Bereich des Gesichtes wurden zwei Elektroden als EOG zur Messung der Augenbewegungen und drei Elektroden rechts und links auf dem Kinn sowie submental als EMG zur Messung des Muskeltonus und der Bewegungen befestigt. Für die Ableitung des EKG wurden eine Elektroden rechts im 2. Interkostalraum medioklavikulär und eine weitere Elektrode links submammilär angebracht. Ein Bauchgurt mit Lagesensor zeigte die Lage des Schlafenden sowie Lagewechsel an.

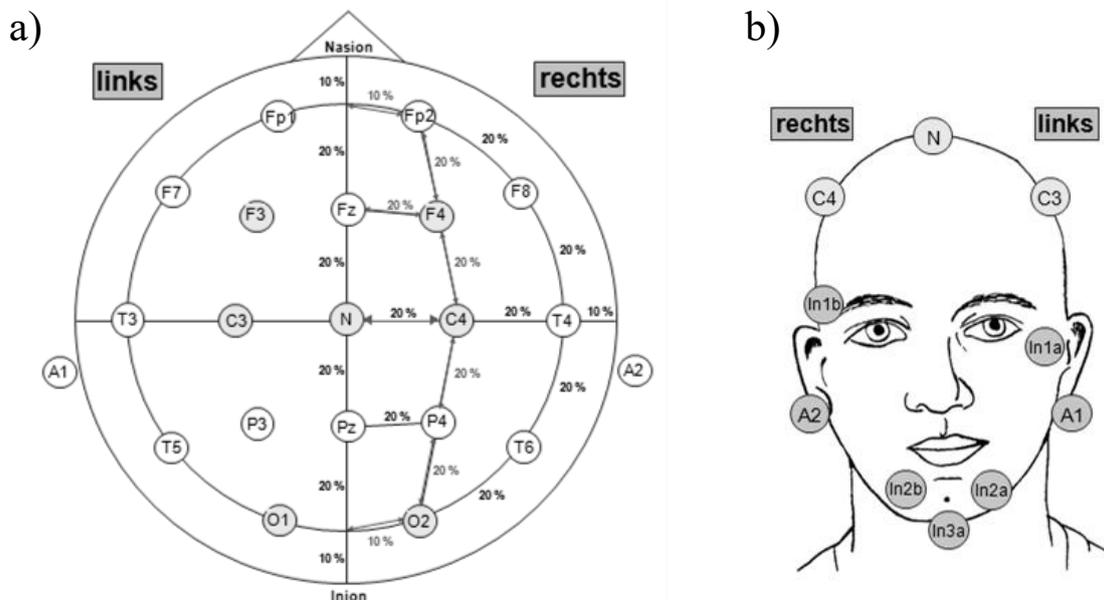


Abb. 10: Position der einzelnen Elektroden. Ansicht von a) kranial und b) frontal. Hauseigene Darstellung des Max-Planck-Instituts für Psychiatrie, München.

3.4.2 Auswertung der Schlafstadien und Spindeln

Die Auswertung der Schlafstadien wurde von erfahrenen Scorerern des MPI für Psychiatrie anhand der AASM-Kriterien vorgenommen (Iber et al. 2007). Sie waren unabhängig und kannten das Studienprotokoll nicht. Im Rahmen der visuellen Auswertung wurde jeder Epoche von 30 Sekunden ein Schlafstadium zugewiesen. Hierfür wurden die EEG-Ableitungen F3-A2, F4-A1, C3-A2, C4-A1, O1-A2 und O2-A1 verwendet und neben der Frequenz und der Amplitude des EEGs weitere stadien-typische Faktoren berücksichtigt. Für die Merkmale der einzelnen Stadien s. 1.2.1.

Die Anzahl der Schlafspindeln wurde mithilfe eines im Haus entwickelten automatischen Algorithmus ermittelt (Supplement zu Genzel et al. 2015a): Hierfür wurden in einem ersten Schritt hochfrequente Muskelartefakte identifiziert und von der Spindelanalyse ausgeschlossen. Auch wurden Segmente mit starker Alpha-Frequenz durch den Logarithmus identifiziert und von der Analyse ausgeschlossen, da Anteile daraus zu falsch positiven Ergebnissen führen können. Die Amplitude des EEG-Signals kann zwischen Versuchspersonen und Kanälen variieren, deshalb berechnete der Algorithmus die Schwellenwerte Sigma-Aktivität (SA) und Sigma-Peak (SP) anhand einer Formel. Für die eigentliche Erkennung der Spindeln wurde die "Continuous Wavelet Transform"-Methode verwendet (Addison 2002). Diese Methode bietet eine gute zeitliche und räumliche Auflösung und liefert Wavelets, die Schlafspindeln ähneln. Das verwendete Wavelet entsprach einem Frequenzbereich von 12-16 Hz. Eine Schlafspindel wurde als solche erkannt, wenn das Ergebnis der CWT den Schwellenwert SA für einen Zeitabschnitt von mindestens 0,5 Sekunden überschritt und mindestens einmal einen Wert oberhalb des Schwellenwertes SP erreichte.

3.5 Statistik

Zur inferenzstatistischen Analyse der Einflüsse von Schlaf und wiederholtem Testen wurde eine ANOVA mit Messwiederholung durchgeführt. Hierbei wurden die Innersubjektfaktoren Zeit (3 Stufen: Test 1, Test 2, Test 3) und Nap (2 Stufen: Nap, Wach) sowie der Zwischensubjektfaktor Bedingung (2 Stufen: wiederholtes Testen, wiederholtes Lernen) verwendet. Dabei wurde bei Bedarf nach Greenhouse-Geisser (1959) korrigiert. Für signifikante Ergebnisse wurden anschließend Innersubjekt-kontraste

zur Aufdeckung der konkreten Richtung der gefundenen Unterschiede getestet. Für alle Analysen wurde ein α von $p < .05$ als signifikant gewertet.

Zur Testung des Zusammenhangs der Variablen N2, N3, REM, Spindelanzahl und Spindeldichte mit den relativen Erinnerungsraten am Testtag und verglichen mit dem Retest nach 4 Tagen wurden jeweils partielle Korrelationen mit der Kontrollvariablen „Bedingung“ berechnet. Zur Testung der Frage, ob lernbedingte Veränderungen des Schlafes mit schlafbedingten Verbesserungen der Gedächtnisleistung zusammenhängen wurde getestet, ob die Abweichungen der Schlafvariablen im Mittagsschlaf am Testtag vom Kontroll-Mittagsschlaf mit den Abweichungen der Vergessensraten in der Schlaf-Bedingung von der Wach-Bedingung korrelieren. Für alle Korrelationsanalysen wurde ein Bonferroni-korrigiertes α von $p < .05$ als signifikant bewertet.

Um festzustellen, ob die Ergebnisse durch Abweichungen der Vigilanz zwischen Wach- und Nap- Tag beeinflusst wurden, wurde eine ANOVA mit Messwiederholung für die Innersubjektvariable Nap (2 Stufen: Nap vs. Wach) und der Zwischensubjektvariablen Bedingung (2 Stufen: Lernen vs. Testing) gerechnet. Auch um mögliche Veränderungen zwischen Nap und Kontrollnap in Bezug auf Schlafparameter festzustellen, wurde für die Schlafstadien eine MANOVA und für die Anzahl der Schlafspindeln und die Spindeldichte jeweils eine ANOVA mit Messwiederholung für die Innersubjektvariable Nap (2 Stufen: Nap vs. Kontrollnap) und die Zwischensubjektvariable Bedingung (2 Stufen: Lernen vs. Testing) gerechnet.

4 Ergebnisse

4.1 Schlafauswertung

Die Probanden hatten eine durchschnittliche Gesamtschlafzeit von 69,0 Minuten und verbrachten davon im Schnitt 22,0 Minuten im Stadium N1, 29,0 Minuten im Stadium N2 und 13,1 Minuten im Stadium N3 (Tiefschlaf). Am kürzesten dauerte die REM-Phase mit durchschnittlich 4,9 Minuten. Somit machte der Leichtschlaf den weitaus größeren Teil des Schlafes aus, verglichen mit Tiefschlaf und REM. Beim Nap erreichten 19 der 80 Probanden nicht das Stadium N3 und 34 Probanden nicht das Stadium REM. Beim Kontrollnap waren es 19 ohne Tiefschlaf und 35 ohne REM. Die genaue Aufschlüsselung der Strukturparameter nach Testbedingung und Testtag mit Mittelwert und Standardabweichung ist der untenstehenden Tabelle zu entnehmen.

4.1.1 Strukturparameter

	Testbedingung				Lernbedingung			
	Nap		Kontrollnap		Nap		Kontrollnap	
	M	St.Abw.	M	St.Abw.	M	St.Abw.	M	St.Abw.
N1	21,6	14,8	21,1	14,6	20,5	11,2	24,8	13,5
N2	27,2	15,5	27,1	14,7	31,3	13,9	30,5	16,9
N3	14,9	15,4	11,9	13,2	14,9	14,9	11,0	10,6
NREM	63,6	21,0	60,0	20,6	66,8	22,1	66,3	16,1
REM	4,4	6,7	4,8	5,8	4,9	6,7	5,4	7,0

Tabelle 1: Mittelwerte (M) und Standardabweichungen (St.Abw.) der verschiedenen Schlafstadien (N1-N3) sowie des gesamten Non-REM und REM-Schlafes, aufgegliedert nach Bedingung und Testtag [min]

Rein deskriptiv fallen für die Gesamtstichprobe der Probanden Unterschiede im Tiefschlaf zwischen Nap und Kontrollnap auf: Im Nap hatten sie durchschnittlich $14,9 \pm 15,0$ Minuten Tiefschlaf, und im Kontrollnap durchschnittlich $11,4 \pm 11,9$ Minuten. Die MANOVA der Schlafparameter zeigte jedoch keinen signifikanten Effekt der Innersubjektvariablen Nap ($F_{4,75}=1,13$, $p=.35$) oder der Zwischensubjektvariablen Bedingung ($F_{4,75}=0,55$, $p=.70$) auf die Schlafparameter an. Weiterhin lässt sich keine signifikante Interaktion der beiden Variablen feststellen ($F_{4,75}=0,87$, $p=.49$). Die Naps in der Experimentalbedingung unterschieden sich somit nicht von den Kontrollnaps, und Lern- und Test-Bedingung hatten keine differenziellen Effekte auf die Naps.

4.1.2 Spindelanalyse

	Nap	Kontrollnap	df	F	p
Spindelanzahl	$228,2 \pm 152,0$	$218,6 \pm 172,9$	1,78	0,43	0,52
Spindeldichte	$5,4 \pm 2,9$	$5,5 \pm 2,8$	1,74	0,19	0,66

Tabelle 2: Spindelanzahl und Spindeldichte (durchschnittliche Anzahl der Spindeln pro gewertete Epoche á 30 Sekunden), aufgeschlüsselt nach Testtag. Jeweils mit Standardabweichung. Angegeben sind weiterhin Freiheitsgrade (df), sowie F-Wert und Signifikanz einer ANOVA mit Messwiederholung zum Vergleich der Spindelanzahl (SP) und Spindeldichte (SPD) zwischen Nap und Kontrollnap.

Anhand der von Artefakten bereinigten EEGs wurde eine Analyse der Spindeln durchgeführt (s. Tabelle 2). Hierbei konnte weder für die Gesamtzahl der Spindeln im Schlaf, noch für die Spindeldichte ein signifikanter Unterschied zwischen dem Nap nach Lernen und dem Kontrollnap gefunden werden. Weiterhin konnte keine signifikante Interaktion zwischen Nap und Testbedingung gefunden werden, Lern- und Test-Bedingung hatten somit keine differenziellen Effekte auf die Spindelparameter.

4.2 Testauswertung

4.2.1 SSS und PVT

	1	2	3	4	5	6	7
SSS_Nap	7	32	26	11	4	0	0
SSS_Wach	4	23	23	23	7	0	0

Tabelle 3: Selbsteinschätzung der Probanden in Bezug auf ihre Wachheit mit Hilfe der Stanford Sleepiness Scale. Werte von 1 (hellwach) bis 7 (am Einschlafen). Vergleich zwischen Nap- und Wachbedingung.

	gesamt				Testing		Lernen	
	nap	wach	F _{1,78}	p	nap	wach	nap	wach
SSS	2.66±0.99	3.08±1.07	6.75	.01	2,68 ± 1,05	3,05 ± 0,98	2,65 ± 0,98	3,10 ± 1,15
PVT	1.59±2.62	1.60±1.93	0.01	.97	1,48 ± 1,55	1,50 ± 1,92	1,70 ± 3,38	1,70 ± 1,96

Tabelle 4: Ergebnisse der Stanford Sleepiness Scale (SSS) und Anzahl der Male, wo die Reaktionsgeschwindigkeit beim PVT mehr als 500ms beträgt. Angabe der Werte als Mittelwert ± Standardabweichung, aufgeschlüsselt nach Testtag (links, mit Ergebnissen der ANOVA) und nach Bedingung und Testtag (rechts).

Die ANOVA mit Testwiederholung zeigte einen signifikanten Effekt der Nap-Bedingung auf die Stanford Sleepiness Scale an: Die Probanden fühlten sich in ihrer Selbsteinschätzung am Testtag nach dem Nap statistisch signifikant wacher als an dem Testtag, an dem sie vorher den Film angesehen hatten. Die genaue Verteilung der Probanden auf die Items des Fragebogens ist Tabelle 3 zu entnehmen. In beiden Fällen waren sie morgens um 7 Uhr aufgestanden und füllten nachmittags den Fragebogen aus. Wie zu erwarten konnten sie durch einen Nap das Nachmittagstief ein wenig abfangen. Der Mittelwert liegt bei beiden Testtagen im Bereich von Item 3 (Entspannt; wach; nicht voll aufmerksam; ansprechbar). Die subjektiven Vigilanzunterschiede konnten jedoch nicht durch objektive Daten im Psychomotor Vigilance Task bestätigt werden, hier gab es deutlich keinen signifikanten Unterschied bei der Anzahl der Versuche mit einer Reaktionszeit ≥ 500 ms. Der PVT gilt dabei als Goldstandard der Vigilanzmessung. Eine

genauere Aufschlüsselung der Ergebnisse nach Bedingung findet sich in Tabelle 4. Werden die Werte für SSS und PVT der Probanden der Test- vs. der Lernbedingung mithilfe einer ANOVA verglichen, ergeben sich keine signifikanten Interaktionseffekte zwischen Wach- und Test-Bedingung ($F_{2,77}=0.03$, $p=.97$).

4.2.2 IQ-Test

Der mithilfe des CFT ermittelte Intelligenzquotient der Probanden liegt mit einem Mittelwert von 115,5 (Median 117) wie für ein rein studentisches Probandenkollektiv zu erwarten über dem allgemeinen Bevölkerungsdurchschnitt, der als 100 definiert wird. Die Einzelwerte liegen zwischen minimal 76 und maximal 142. Test- und Lernbedingung unterscheiden sich lediglich leicht in der Varianz, während Mittelwert und Median beider Bedingungen identisch sind (s. Tabelle 5). Somit ist ein Confounding-Effekt auf die Gedächtnisaufgabe oder andere Variablen wie die Spindelaktivität durch den IQ-Wert ausgeschlossen.

	N	Mittelwert	Median	St.Abw.	Min	Max
Gesamt	80	115,5	117,0	16,3	76	142
Test	40	115,5	117,0	14,7	76	142
Lern	40	115,5	117,0	17,8	80	142

Tabelle 5: IQ-Werte des Gesamtprobandenkollektivs sowie aufgeschlüsselt nach Bedingungen unter Angabe der jeweiligen Probandenanzahl, des jeweiligen Mittelwerts, Medians, Standardabweichung sowie des Minimal-und Maximalwertes innerhalb der Gruppe.

4.2.3 Lernaufgabe

Nach der Lernaufgabe konnten die Probanden im Schnitt 22,6 der insgesamt 36 Wortpaare im Rahmen des ersten Tests (T1, Baseline) vor dem Nap bzw. Film richtig wiedergeben. Die Ausgangswerte sind für die Probanden der Lern- und der Testbedingung nahezu identisch (22,8 vs. 22,4), so dass sie wie im Rahmen des Versuchsdesigns angestrebt trotz verschiedener Formen des Wissenserwerbs ein gleiches Ausgangsniveau haben. Mit einer durchschnittlichen Quote von 63% richtig erinnerter

Wortpaare liegen sie etwas unter den angestrebten 70%, aber es ist deutlich, dass für die Mehrheit der Probanden Decken- und Bodeneffekte vermieden werden konnten.

	Testbedingung				Lernbedingung			
	Nap		Wach		Nap		Wach	
	M	St.Abw.	M	St.Abw.	M	St.Abw.	M	St.Abw.
T1	22,2	7,7	22,6	8,0	22,1	9,0	23,5	8,1
T2	19,6	7,4	19,0	8,1	18,9	9,7	18,7	8,3
T3 (re)	17,4	7,7	17,5	8,2	14,9	8,9	16,4	9,1
T2/T1	88,1%	12,0%	82,4%	12,7%	82,4%	16,8%	78,3%	15,3%
T3/T1	78,1%	18,7%	74,7%	17,9%	66,1%	21,7%	66,6%	23,2%
1-T2/T1	11,9%	12,0%	17,6%	12,7%	17,6%	16,8%	21,8%	15,3%
1-T3/T1	22,0%	18,7%	25,3%	17,9%	33,9%	21,7%	33,4%	23,3%

Tabelle 6: Mittelwerte (M) und Standardabweichungen (St.Abw.) der Testergebnisse in der Lernaufgabe vor dem Nap/Film (T1), danach (T2) und beim Retest nach 4 Tagen (T3 (re)). Außerdem die Behaltensrate zwischen T1 und T2 (T2/T1) sowie zwischen T1 und T3 (T3/T1), sowie die entsprechenden Vergessensraten (1-Behaltensrate). Alle Daten aufgeschlüsselt nach Bedingung und Testtag.

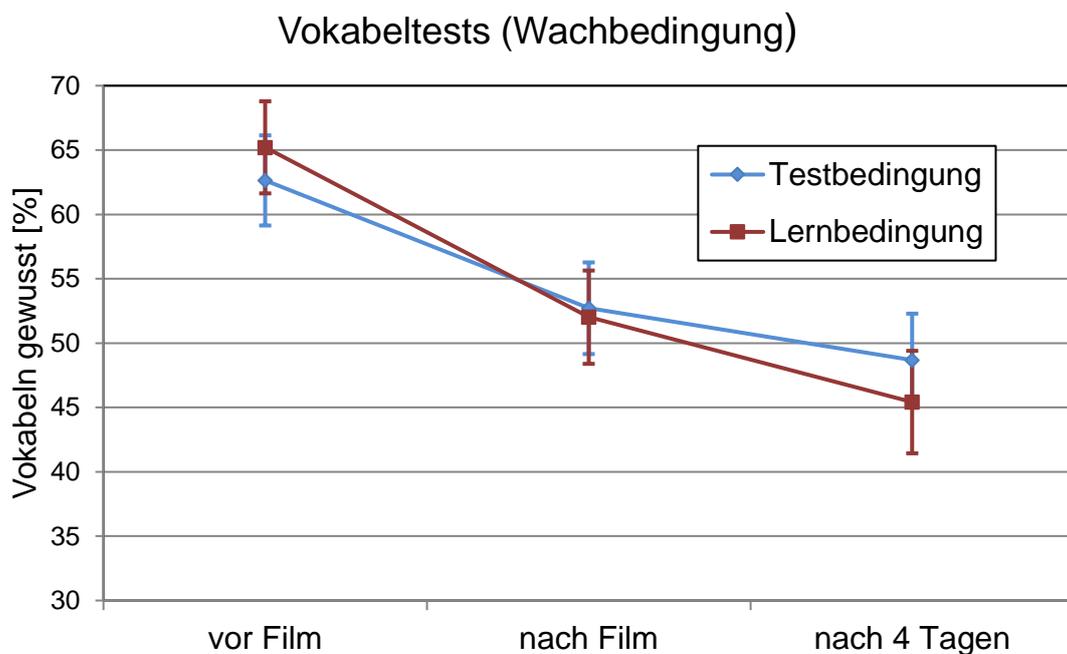
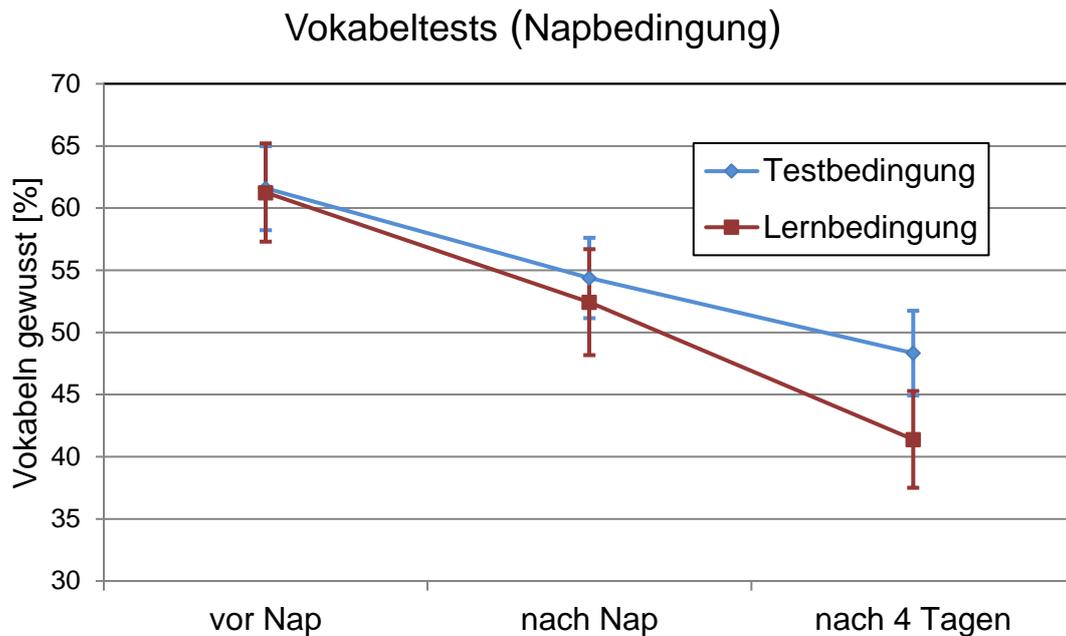


Abb. 11: Anteil der korrekt erinnerten Wortpaare im zeitlichen Verlauf (T_i / N ; $N=36$). Daraus ergeben sich die Vergessenskurven für die Napbedingung (oben) und Wachbedingung (unten). Es werden jeweils Probanden der Test- und der Lernbedingung verglichen. Fehlerbalken zeigt Standardabweichung des Mittelwerts.

Wie erwartet nimmt die Behaltensleistung von Test zu Test ab. Betrachtet man die Entwicklung der Mittelwerte über Zeit (s. Abb. 11), so lässt sich qualitativ feststellen, dass die Vergessenskurve für die Lernbedingung allgemein steiler verläuft, als für die Testbedingung, d.h. dass Probanden der Lernbedingung tendenziell schneller vergessen.

Bei der Wachbedingung verläuft die Vergessenskurve zwischen erstem und zweitem Test für beide Gruppen steiler als bei der Nap-Bedingung. Graphisch etwas anders aufbereitet finden sich diese Informationen auch im Vergleich der relativen Vergessensraten. Vergleicht man die Vergessensraten der Probanden der Test- und Lernbedingung (gemittelt über beide Testtage), so zeigen die Probanden der Testbedingung sowohl beim Test am gleichen Tag, als auch beim Test nach vier Tagen geringere Vergessensraten (s. Abb. 12). Vergleicht man hingegen bei allen Probanden (Lern- und Testbedingung zusammen) die Vergessensraten zwischen Nap- Bedingung und Wachbedingung, so zeigt sich ein Unterschied am Tag selbst, während vier Tage später kaum noch ein Unterschied besteht (s. Abb. 13). Welche dieser qualitativ sichtbaren Unterschiede auch statistisch signifikant sind, ist der inferentiellen Statistik in Abschnitt 4.3.2 zu entnehmen.

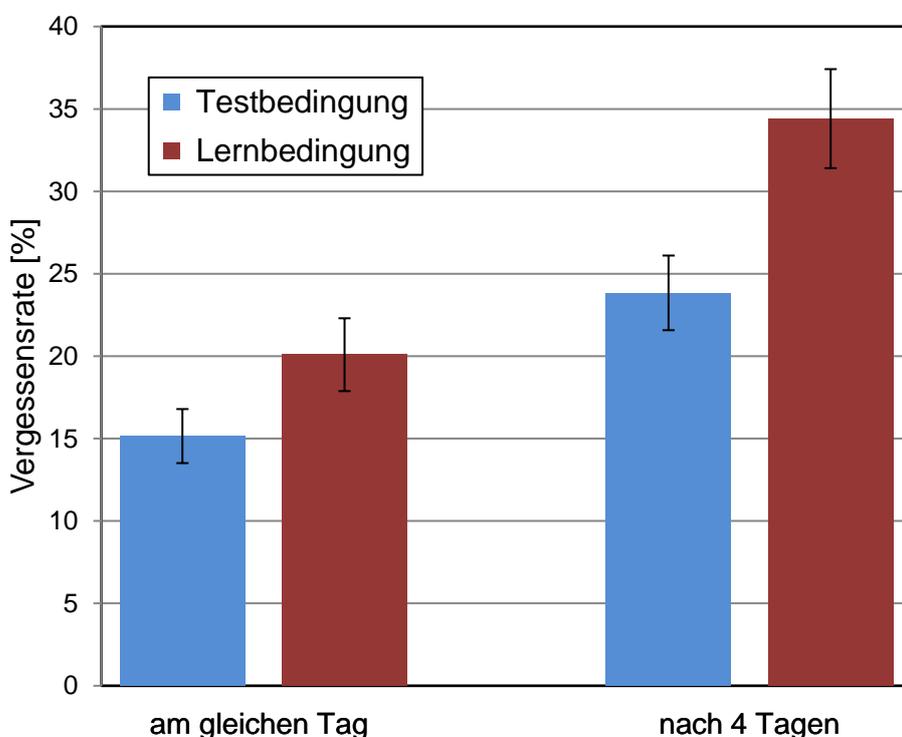


Abb. 12: Relative Vergessensrate am gleichen Tag ($1-T_2/T_1$) und nach 4 Tagen ($1-T_3/T_1$). Vergleich der Test- und Lernbedingung. Fehlerbalken zeigen Standardfehler des Mittelwerts.

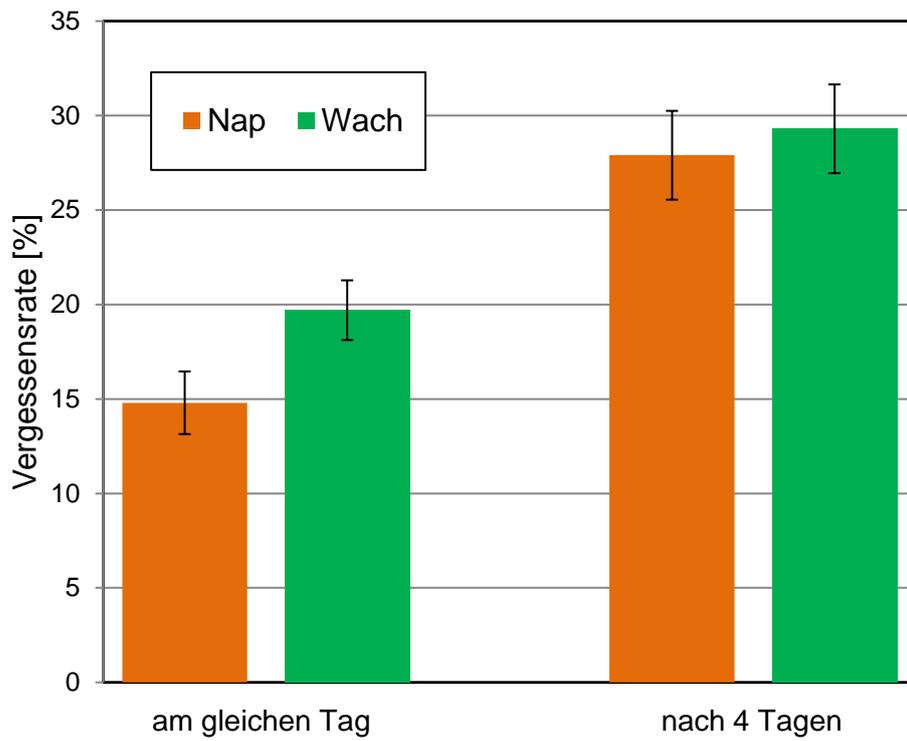


Abb. 13: Relative Vergessensraten am gleichen Tag ($1-T_2/T_1$) und nach 4 Tagen ($1-T_3/T_1$). Vergleich der Nap- und Wachbedingung. Fehlerbalken zeigen Standardfehler des Mittelwerts.

4.3 Inferentielle Statistik

4.3.1 Schlafvariablen und Gedächtnisleistung

	Erinnerungsrate Testtag	Erinnerungsrate Retest
N2	$r=-.09, p=.46$	$r=-.11, p=.37$
N3	$r=.30, p=.01$	$r=-.10, p=.37$
REM	$r=-.03, p=.77$	$r=-.09, p=.43$
Spindelanzahl	$r=-.07, p=.53$	$r=-.03, p=.77$
Spindeldichte	$r=-.23, p=.05$	$r=.04, p=.75$

Tabelle 7: Partielle Korrelation der einzelnen Schlafparameter (Stadium N2, N3 und REM, sowie Spindelanzahl (SP) und Spindeldichte (spd) mit den relativen Erinnerungsraten am Testtag direkt nach dem Nap bzw. bei dem Retest (T3) 4 Tage nach dem Nap. Relevante Ergebnisse fettgedruckt, Kommentar s. Text.

Weder für den zweiten Test unmittelbar nach dem Mittagsschlaf, noch für den dritten Test nach vier Tagen (Retest) konnten nach Bonferroni-Korrektur signifikante Korrelationen der Schlafvariablen N2, N3, REM, Spindelanzahl oder Spindeldichte mit der Gedächtnisleistung festgestellt werden (vgl. Tabelle 7). Ohne Bonferroni-Korrektur korrelierte N3 positiv und die Spindeldichte negativ mit der Gedächtnisleistung am Testtag. Wurde ausschließlich die Behaltensrate am Testtag selbst betrachtet, so korrelierte N3 auch mit Bonferroni-Korrektur positiv, d.h. mehr Tiefschlaf führte zu einem größeren Prozentsatz erinnerter Wortpaare. Für eine analoge Berechnung der Korrelationen zwischen den spezifischen Abweichungen des Mittagsschlafs am Testtag vom Kontroll-Mittagsschlaf ohne Lerneinflüsse konnten – mit oder ohne Bonferroni-Korrektur – keine signifikanten Korrelationen mit den entsprechenden differenziellen Gedächtnisdaten beobachtet werden (alle $|r|<.17$, alle $p>.15$).

4.3.2 Effekte von Schlaf und wiederholtem Testen

Die ANOVA mit Messwiederholung identifizierte einen signifikanten Effekt des Faktors Zeit ($F_{2,77}=150.7, p<.001$), eine signifikante Interaktion der Faktoren Zeit und Nap ($F_{2,77}=8.4, p=.001$) und eine signifikante Interaktion der Faktoren Zeit und Bedingung

($F_{2,77}=4.3$, $p=.017$). Es bestanden hingegen kein allgemeiner Effekt des Faktors Nap ($F_{1,78}=0.9$, $p=.342$) und keine Interaktionen Nap \times Bedingung ($F_{1,78}=1.1$, $p=.307$) oder Zeit \times Nap \times Bedingung ($F_{2,77}=0.6$, $p=.531$).

Source	nap	time	Sum of Squares	df	Mean Square	F	Sig.
nap	2 vs. 1		15.313	1	15.313	.912	.342
nap * cond	2 vs. 1		17.735	1	17.735	1.057	.307
Error(nap)	2 vs. 1		1309.286	78	16.786		
time		2 vs. 1	994.050	1	994.050	210.331	.000
		3 vs. 1	2898.028	1	2898.028	250.858	.000
time * cond		2 vs. 1	15.313	1	15.313	3.240	.076
		3 vs. 1	100.128	1	100.128	8.667	.004
Error(time)		2 vs. 1	368.637	78	4.726		
		3 vs. 1	901.094	78	11.552		
nap * time	2 vs. 1	2 vs. 1	130.050	1	130.050	13.881	.000
		3 vs. 1	1.012	1	1.012	.037	.849
nap * time * cond	2 vs. 1	2 vs. 1	7.200	1	7.200	.769	.383
		3 vs. 1	1.513	1	1.513	.055	.816
Error(nap*time)	2 vs. 1	2 vs. 1	730.750	78	9.369		
		3 vs. 1	2154.475	78	27.621		

Tabelle 8: Überblick über die Innersubjektivitätskontraste. F-Werte und Signifikanz;
In Spalte time: 1= Baseline, Test vor Nap/Film; 2= Test, Test nach Nap/Film; 3= Retest,
Test nach 4 Tagen; Weitere Interpretation s. Text

Anschließende Tests der Innersubjektkontraste (s. Tab. 8) zeigten eine signifikante Interaktion zwischen Zeit und Nap für den Kontrast zwischen Test (2) und Baseline (1) ($F=13.9$, $p<.001$), nicht jedoch zwischen Retest nach vier Tagen (3) und Baseline (1) ($F=.04$, $p=.849$). Umgekehrt konnte zwar keine signifikante Interaktion zwischen Zeit und Bedingung für den Kontrast zwischen Test und Baseline ($F=3.2$, $p=.076$), jedoch schon zwischen Retest und Baseline ($F=8.7$, $p=.004$) gefunden werden. Probanden der Test-Bedingung unterscheiden sich somit noch nicht beim Test am Versuchstag (2), wohl aber im Retest nach vier Tagen (3) signifikant in ihrer Vergessensrate von den Probanden der Lern-Bedingung. Weiterhin unterscheiden sich Probanden in der Nap-Bedingung signifikant am Versuchstag, jedoch nicht mehr im Retest nach vier Tagen in ihrer Vergessensrate von den Probanden der Wach-Bedingung.

5 Diskussion

5.1 Zusammenfassung der Ergebnisse

Probanden der Lern- bzw. Testbedingung zeigten keine signifikanten Unterschiede der Schlafparameter, ebenso fanden sich keine signifikanten Abweichungen der Parameter zwischen dem Nap am Testtag und dem Kontroll-Nap. Die Varianzanalyse zeigte signifikante Werte für die Faktoren Zeit, Zeit x Nap, sowie Zeit x Bedingung. Allerdings zeigte sich keine Interaktion Zeit x Nap x Bedingung. Bei den anschließenden Tests der Innersubjektivitätskontraste zeigte sich für die Interaktion Zeit x Nap ein signifikanter Unterschied ausschließlich zwischen Test und Baseline, für die Interaktion Zeit x Bedingung hingegen ausschließlich zwischen Retest und Baseline. Folglich war am Testtag selbst ein signifikanter Unterschied zwischen Probanden der Nap-Bedingung und Probanden der Wachbedingung, der nach vier Tagen nicht mehr sichtbar war. Umgekehrt war zwischen Probanden der Testbedingung und denen der Lernbedingung am Testtag selbst kein signifikanter Unterschied, wohl aber beim Retest nach vier Tagen. Ein Einfluss von Schlaf auf den Testing Effekt fand sich nicht. Tiefschlaf korrelierte positiv mit der relativen Behaltensrate am Testtag. Spindeln zeigten keinerlei Korrelationen mit kognitiver Leistung.

5.2 Diskussion der Ergebnisse

Die Ergebnisse der Studie werden im Folgenden ausgehend von den in der Zielsetzung formulierten Fragestellungen diskutiert.

5.2.1 Effekte von Schlaf und wiederholtem Testen auf die Erinnerungsleistung

Im Rahmen der Studie sollte herausgefunden werden, ob die vielfach vorbeschriebenen positiven Effekte von Schlaf sowie die erst verzögert auftretenden positiven Effekte des wiederholten Abrufes auf die Erinnerungsleistung mithilfe eines großen Probandenkollektivs bestätigt und eine potentielle Interaktion gezeigt werden kann. Wie erwartet haben die teilnehmenden Probanden am Tag der Nap-Bedingung signifikant geringere Vergessensraten als am Tag der Wachbedingung. Die subjektiv höhere Vigilanz der Probanden nach Nap (gemäß SSS) könnte hierbei als Confounding-Faktor fungieren.

Allerdings bestätigt sich dieser Unterschied der Vigilanz nicht im PVT, der als objektiver Parameter und als Goldstandard der Vigilanzmessung gilt. Eine Interferenz durch die Inhalte des Films (siehe auch 5.4.) könnte die Vergessensraten im Rahmen der Wachbedingung ebenfalls etwas mit beeinflussen. Allerdings kann, nicht zuletzt aufgrund der breiten Literatur hierzu (Jenkins und Dallenbach 1924; Plihal und Born 1997; Gais 2006; Ficca et al. 2010), davon ausgegangen werden, dass der Schlaf als solcher hier die wichtigste Rolle spielt. Allerdings verschwindet der Nap-Effekt im Lauf der Zeit, so dass bei einem weiteren Test nach vier Tagen kein Unterschied zwischen den beiden Bedingungen mehr zu erkennen ist. Ein Erklärungsansatz hierfür ist, dass auch die Probanden der Wachbedingung in den folgenden Nächten die erlernten Wortpaare im Schlaf konsolidiert haben. Möglicherweise profitieren sie relativ stärker als die Probanden der Nap-Bedingung, die bereits konsolidiert haben, von dem folgenden Nachtschlaf. So holen die Probanden der Wachbedingung durch die Konsolidierung während des Schlafes der folgenden Nächte den ursprünglichen Abstand zu der Nap-Bedingung wieder auf und der positive Nap-Effekt verschwindet sukzessive.

Erwartungsgemäß hatten die Probanden der Testbedingung, die die Wortpaare mithilfe wiederholten Abrufes geübt hatten, geringere Vergessensraten als die Probanden der Lernbedingung. Während sich diese Tendenz schon am ersten Tag andeutet (s. Tabelle 6 sowie Abb. 11 und 12), wird der Unterschied zwischen den beiden Gruppen erst bei dem erneuten Test nach vier Tagen statistisch signifikant. Dieses Ergebnis steht im Einklang mit zahlreichen anderen Studien, die ebenfalls zu dem Ergebnis kamen, dass der Testing-Effekt erst mit einer gewissen Zeitverzögerung auftritt (Wheeler et al. 2003; Roediger und Karpicke 2006a; Toppino und Cohen 2009; Congleton und Rajaram 2012; van den Broek et al. 2014). In der letztgenannten Studie fanden sich schon unmittelbar nach der Lernphase kürzere Antwortzeiten für wiederholt abgerufene verglichen mit wiederholt gezeigten Items, aber erst mit zeitlicher Verzögerung zeigte sich ein positiver Testing Effekt für die Anzahl der korrekt erinnerten Items. In beiden Bedingungen verlangsamten sich die Antwortzeiten über die Zeit in vergleichbarem Maße. Für Items die in der Übungsphase mehrfach korrekt abgerufen wurden, konnte auch ohne zeitliche Verzögerung ein Testing Effekt gezeigt werden (van den Broek et al. 2014). Die Frage, warum der Testing Effekt scheinbar erst zeitlich verzögert auftritt, ist von großer praktischer Relevanz, da sie die Ausgangslage weiterer Studien wie der hier vorliegenden Arbeit darstellt. Darüber hinaus werden die verschiedenen theoretischen Erklärungsansätze des Testing Effekts unter anderem daran gemessen, inwieweit sie

dieses zeitlich verzögerte Auftreten erklären können (Roediger und Karpicke 2006b). Unter 5.2.2 werden die verschiedenen theoretischen Ansätze im Licht der vorliegenden Ergebnisse diskutiert.

Zusammenfassend konnten die beiden Grundannahmen, auf denen die Hypothese der Arbeit fußt, nämlich dass sowohl Schlaf als auch wiederholter Abruf eine positive Wirkung auf die Gedächtnisleistung haben, im Rahmen der Studie bestätigt werden.

5.2.2 Interaktion von Schlaf und Testing Effekt sowie alternative

Erklärungsmodelle

Der Ausgangspunkt der vorliegenden Arbeit war die Frage, ob das verzögerte Auftreten des Testing Effekts durch eine Interaktion von Schlaf und Testing begründet ist. Unter anderem für künftig relevante Inhalte (Wilhelm et al. 2011) konnte bereits eine derartige positive Interaktion mit Schlaf im Sinne einer „Triage“ von Gedächtnisinhalten gezeigt werden (vgl. Stickgold und Walker 2013). Werden also auch getestete Items anders im Schlaf konsolidiert als mehrfach gelernte? Im Rahmen der vorliegenden Arbeit zeigt sich für die Interaktion von Schlaf und Testing keinerlei statistische Signifikanz, so dass sich die Annahme einer solchen Interaktion nicht bestätigte. Wenngleich schwache modulierende Effekte des Schlafes nicht ausgeschlossen sind, so scheint zumindest keine starke Interaktion i.S. einer grundlegend anderen Verarbeitung vorzuliegen, denn von dieser wäre zu erwarten gewesen, dass sie nach dem Nap sichtbar wird. Konfundierende Interaktionen der Vigilanz mit der Testbedingung konnten ebenfalls ausgeschlossen werden. Es gibt bisher kaum Literatur zu dieser Frage, besondere Aufmerksamkeit verdient allerdings eine kürzlich erschienene Arbeit von Bäuml und Kollegen (2014). In einer Reihe von Experimenten (wobei Art des erlernten Materials, Schwierigkeit des Abrufs durch Präsentation von Interferenz-Items, und Anzahl der Wiederholungen in der Übungsphase variiert wurden) kamen sie übereinstimmend zu dem Ergebnis, dass Schlaf den Testing Effekt sogar verringern kann. Dieser Effekt beruht vermutlich darauf, dass Schlaf die Erinnerung an wiederholt gelernte Items fördert, während die durch Testing geübten Items nicht zusätzlich von dem Schlaf profitieren. Bei kurzem Intervall (12 Minuten) zeigte sich kein Testing Effekt, während bei einem 12h-Wachintervall ein deutlicher Testing Effekt auftrat. Bei einem 12h-Intervall mit Nachtschlaf hingegen war der Testing Effekt verglichen mit dem Wachintervall je nach Experiment deutlich abgeschwächt oder gänzlich aufgehoben. Weitere statistische Analysen offenbarten, dass wiederholt gelernte Items (Lernbedingung) von Schlaf profitierten, während die

getesteten Items (Testbedingung) keinen solchen positiven Schlafeffekt zeigten. Beim Vergleich der Ergebnisse von Bäuml und Kollegen mit der vorliegenden Studie sind einschränkend verschiedene methodische Unterschiede zu beachten: Bäuml und Kollegen verwendeten weniger und rein deutsche Wortpaare, lernten mit weniger Wiederholungen, hatten bei ihren Experimenten mehrheitlich weibliche Versuchspersonen, das Testing fand ohne Feedback statt, anstelle des Nap wurde eine volle Nacht geschlafen und es wurde lediglich ein Test 12 Stunden nach der Lernphase durchgeführt und ausgewertet. Sieht man über diese methodischen Unterschiede hinweg und vereint die Ergebnisse ihrer Arbeit mit denen der hier vorliegenden Studie so kommt man zu dem Ergebnis, dass 1) unmittelbar nach der Lernphase noch kein Testing Effekt besteht, 2) nach einem kurzen Intervall von 2-3 Stunden mit oder ohne Nap ein positiver Effekt von Schlaf aber kein Effekt von Testing besteht 3) nach einem 12-stündigen Intervall mit oder ohne Schlaf ein positiver Schlaf- und ein positiver Testing Effekt besteht, wobei der Testing Effekt bei Schlaf geringer ist als bei einem Wachintervall oder ganz verschwindet und 4) nach vier Tagen nur noch der Testing Effekt, nicht aber der Schlafeffekt sichtbar ist. Ein solcher zeitlicher Verlauf wäre mit der Bifurkationshypothese vereinbar. Da die erfolgreich abgerufenen Items deutlich verstärkt werden, während die erneut gelernten gleichmäßig relativ schwächer verstärkt werden, bleiben die erfolgreich abgerufenen Items auch nach einem längeren Zeitintervall über der Abrufbarkeitsschwelle, so dass sie nicht zusätzlich von dem Schlaf profitieren (Bäuml et al. 2014). Im Rahmen der Bifurkationshypothese ist der zeitliche Verlauf auch bei gleichem Effekt von Schlaf auf gelernte und getestete Wortpaare denkbar. Eine grundlegend verschiedene Prozessierung von gelernten und getesteten Wortpaaren im Schlaf wird durch die Ergebnisse der vorliegenden Arbeit zumindest nicht gestützt. Als alternativen Erklärungsansatz schlagen Bäuml und Kollegen vor, dass gerade die lange Wachphase mit ihren sensorischen Eindrücken, die als retroaktive Interferenz die Abrufbarkeitsschwelle erhöhen, zum Auftreten des Testing Effekts beitragen, während dieser Interferenzeffekt nach einem Schlafintervall ähnlich gering ist wie direkt nach der Lernphase, so dass in beiden letzteren Fällen kein Testing Effekt sichtbar ist.

5.2.3 Schlafstadien, Spindeln und Lernen

Es sollte auch die Frage beantwortet werden, ob es Korrelationen zwischen Schlafdauer, einzelnen Schlafstadien oder Spindeldichte mit dem individuellen Lernerfolg gibt. Lediglich für den Tiefschlaf (N3) konnte eine positive Korrelation mit der individuellen

Behaltensleistung gezeigt werden. Dies ist in Einklang mit mehreren Studien mit Nachthälften-Design, die bessere Behaltensraten für deklarative Lernaufgaben nach der SWS-reichen ersten Nachthälfte verglichen mit der zweiten Nachthälfte fanden (Yaroush et al. 1971; Fowler et al. 1973; Plihal und Born 1997). Einschränkend sei jedoch erwähnt, dass es u.a. eine sehr groß angelegte Studie gibt, die diesen Effekt nicht findet. Dabei wurden den Versuchspersonen Bilder gezeigt und sie mussten sie am nächsten Tag so gut wie möglich beschreiben. Möglicherweise ist aufgrund des etwas anderen Lernmaterials kein Vergleich möglich, möglicherweise überschätzen aber auch die oben genannten Studien aufgrund des kleinen Probandenkollektivs mögliche Korrelationen stark (Ackermann et al. 2015).

In der vorliegenden Studie zeigte sich keinerlei positive Korrelation zwischen den verschiedenen Spindelparametern und dem Lernerfolg. Im Gegensatz zu Gais et al. (2002) fand die vorliegende Arbeit keinen Unterschied in der Spindeldichte zwischen Nap nach der Lernaufgabe und dem Kontrollnap. Im Gegensatz zu Cox et al. (2012) wurden die Spindelparameter nicht nach Schlafstadien getrennt, allerdings erscheint es unwahrscheinlich, dass durch diese Differenzierung die aktuell negative Korrelation zwischen Spindeldichte und Behaltensleistung signifikant positiv werden könnte. Beide zitierten Arbeiten messen Traits, d.h. sie korrelieren die interindividuell unterschiedlichen Spindeldichten mit dem jeweiligen Lernerfolg. Bei näherer Betrachtung scheint allerdings der State (intraindividuelle Unterschiede der Spindelparameter nach Lernen vs. Kontrollschlaf) das bessere Maß in dieser Fragestellung zu sein (Ackermann et al. 2015). Darüber hinaus arbeiten beide Studien mit relativ kleinen Probandenkollektiven und kontrollieren nicht für Geschlechtsunterschiede. In einer großangelegten Studie untersuchten Ackermann und Kollegen (2015) insgesamt 929 junge Versuchspersonen, die in einer Gedächtnisaufgabe Bilder beschreiben mussten, die sie am Tag zuvor gesehen hatten. Es zeigte sich, dass weibliche Probanden signifikant mehr Bilder erinnern konnten und signifikant mehr Schlafspindeln im Schlaf-EEG der Nacht zwischen Enkodierung und Abruf hatten. Daher sahen die Autoren die Notwendigkeit gegeben, für Geschlecht zu kontrollieren. Wurde für Geschlecht und Alter kontrolliert, fanden sich keinerlei signifikante Korrelation zwischen der Spindeldichte und der Gedächtnisleistung. Daher müssen frühere Erkenntnisse kritisch hinterfragt bzw. mit größeren Kollektiven und kontrolliert für Geschlechts- und Alterseffekte repliziert werden. Am Ende ihrer Arbeit merken Ackermann und Kollegen an, dass ihre Ergebnisse im Hinblick auf die nächtliche

Gedächtniskonsolidierung auf eine größere Rolle der intraindividuellen Unterschiede (State) verglichen mit den interindividuellen Unterschieden (Trait) hinweisen.

5.3 Ausblick

Aufgrund der oben genannten methodischen Unterschiede zwischen der Studie von Bäuml und Kollegen und der hier vorliegenden Studie wäre es für ein tieferes Verständnis des zeitlichen Verlaufs des Testing Effekts sowie einer möglichen negativen Interaktion von Schlaf und dem Testing Effekt wünschenswert, diesen zeitlichen Verlauf im Rahmen eines einheitlichen Studienprotokolls zu untersuchen. Dies könnte geschehen, indem bei einem vergleichbaren Versuchsaufbau die Tests in verschiedenen Zeitintervallen absolviert werden (bspw. 2h, 6h, 12h, 1d, 2d, 4d, 7d) oder gleichzeitig die Schlafzeiten manipuliert werden (bspw. Nap vs. Nachtschlaf vs. 12h Wachintervall vs. 48h Schlafentzug). Aufgrund der Implikationen die Erkenntnisse bzgl. des Einflusses von Schlaf und anderer Faktoren auf den Testing Effekt für die Formulierung und Bewertung von theoretischen Erklärungsmodellen haben, kann weitere Forschung in diesem Bereich wichtige Puzzlesteine zu einem tieferen Verständnis liefern.

Sowohl die Erforschung von Schlaf und Gedächtnis als auch die Erforschung des Testing Effekt werden aktuell mit einer hohen Dynamik betrieben und es bleibt abzuwarten, welche neuen Erkenntnisse im Laufe der kommenden Jahre unser Verständnis der zugrundeliegenden Mechanismen weiter vertiefen oder auch teilweise revolutionieren werden.

5.4 Diskussion der Methoden und Durchführung

Da die Probanden den einzelnen Bedingungen nach dem Zufallsprinzip zugeordnet waren, ist davon auszugehen, dass sich interindividuelle Unterschiede wie Schlafqualität, Sprachlernfähigkeiten oder Erinnerungsvermögen recht gleichmäßig verteilt haben und es hierbei zu keinen größeren Verzerrungen gekommen ist. Allerdings handelt es sich bei dem im Rahmen der Studie akquirierten Probandenkollektiv insgesamt um eine selektive Subpopulation (männliche Studenten), die sich in Alter, Geschlecht und Bildungsgrad ähnlich sind, so dass nicht sicher gesagt werden kann, inwieweit sich die Ergebnisse auf die Gesamtbevölkerung verallgemeinern lassen. Diese Einschränkung teilt die vorliegende Arbeit mit vielen anderen Studien, die vornehmlich aus praktischen Gründen

an Studenten durchgeführt werden. Insbesondere bei Frauen ist bekannt, dass ihr Schlaf und auch das Lernvermögen in Abhängigkeit von den zyklischen Hormonschwankungen diversen Veränderungen unterliegt (Genzel et al. 2012). Auch mit zunehmendem Alter verändert sich die Schlafqualität und die gesamte Schlafarchitektur (Crowley 2011).

Im Design der Studie finden sich einzelne Faktoren, die zu einer gewissen Unschärfe der Daten beitragen könnten: Aufgrund der Tatsache, dass die Probanden an allen Testtagen zur gleichen Zeit im Labor sein mussten, die Gesamtdauer der einzelnen Testtage sich aber unterschied, kam es zu gewissen systematischen Abweichungen in der Uhrzeit einzelner Tests. So musste am Tag der Nap- Bedingung allen Probanden ein EEG geklebt werden, so dass sich die weiteren Tests und auch der Nap häufig um eine halbe bis dreiviertel Stunde gegenüber der Wachbedingung verschoben. Auf der anderen Seite fanden am Tag des Kontroll-Naps keine Versuche statt, es mussten lediglich diverse Fragebögen ausgefüllt werden, so dass der Kontroll-Nap tendenziell zeitlich etwas früher stattfand als der Nap. Deshalb sind leichte zirkadiane Effekte möglich. Auch variierte, wieviel Zeit die Probanden brauchten, um sich nach dem Nap fertig zu machen, so dass es zu gewissen Abweichungen in dem Zeitintervall zwischen den Tests vor und nach dem Nap bzw. Film kam. Allerdings betragen diese Abweichungen nie mehr als maximal 30 Minuten und glichen sich über das Gesamtkollektiv der Probanden nahezu vollständig wieder aus. Da die Probanden neben der Lernaufgabe auch eine Kreativitätsaufgabe mit Brainstorming absolvieren mussten, die am Tag des Kontroll-Naps ebenfalls nicht stattfand, können mögliche Veränderungen der Schlafparameter zwischen Nap und Kontrollnap nicht mit abschließender Sicherheit der Lernaufgabe zugeordnet werden. Ein weit verbreitetes Vorgehen bei Napstudien ist, die Probanden der Wachbedingung mit einem Film zu beschäftigen, um die Zeit zu überbrücken und die zeitliche Parallelität zur Nap-Bedingung zu gewährleisten. Indem die Probanden beschäftigt sind, soll u.a. verhindert werden, dass sie die Wartezeit nutzen um die erlernten Wortpaare zu wiederholen. Es wird bewusst ein Film mit niedrigem emotionalen Gehalt gewählt, um eine stärkere Beeinträchtigung des Gelernten durch die Inhalte des Films zu vermeiden. Allerdings kann man im Sinne der *Interferenztheorie des Vergessens* argumentieren, dass eine schlechtere Behaltensleistung bei Probanden der Wachbedingung verglichen mit den Probanden der Napbedingung u.a. auf die Aufnahme von neuen Informationen durch den Film (=Interferenz) zurückzuführen ist.

Bei der Durchführung der Versuche konnte beobachtet werden, dass einzelne Probanden der Lernbedingung versuchten, einen Teil des Wortpaares abzudecken. Dies lässt auf ein

gewisses – ggf. intuitives – Erfahrungswissen bzgl. des Testing Effekts schließen. Um ein Verwischen der Bedingungen zu vermeiden wurde das Abdecken vom Versuchsleiter unterbunden.

Im Rahmen der Auswertungen fiel auf, dass es bei der Programmierung der Testaufgabe zu einem Fehler bei den Feedbackzeiten gekommen war. Eine der beiden parallelisierten Wortlisten hatte eine Feedbackzeit von fünf Sekunden bei Falschantwort, während die andere Wortliste nur drei Sekunden Feedbackzeit hatte. Allerdings waren die ersten Probanden noch nicht betroffen gewesen und eine kurze Zwischenauswertung zeigte zwar eine leichte Tendenz, aber keine signifikante statistische Auswirkung dieses Unterschieds in den Feedbackzeiten. Trotzdem wurden jeweils vier Probanden in Test- und Lernbedingung zusätzlich getestet, wobei bei den Probanden der Testbedingung die Feedbackzeiten nun umgekehrt auf die Wortlisten verteilt wurden. Da die Wortlisten und somit die Fehler in den Feedbackzeiten gleichmäßig über die verschiedenen Bedingungen verteilt waren, sind keine systematischen Effekte des Programmfehlers auf die Ergebnisse zu erwarten.

6 Zusammenfassung

Die Fähigkeit Wissen zu erwerben und erfolgreich wiederzugeben ist von größter Bedeutung in unserer heutigen Bildungsgesellschaft. Die vorliegende Arbeit befasst sich mit dem Einfluss von Schlaf und dem Testing Effekt auf die deklarative Gedächtnisbildung. Testing Effekt beschreibt das Phänomen, dass das (wiederholte) Abprüfen von erlerntem Wissen in Form von Tests nicht nur zur Feststellung des aktuellen Wissensstands geeignet ist, sondern vielmehr selbst zu einer deutlichen Steigerung der Erinnerungsleistung beiträgt. Viele Studien zeigen eine zeitabhängige Entwicklung des Testing Effekts: Kurz nach der Enkodierung werden durch wiederholtes Testen einstudierte Items ähnlich gut erinnert wie durch mehrfache Präsentation erlernte Items, nach einigen Tagen Intervall werden die getesteten Items jedoch deutlich häufiger korrekt erinnert. Daraus resultiert die Annahme, dass Schlaf und die im Schlaf stattfindenden kognitiven Prozesse möglicherweise in entscheidender Weise für den Testing Effekt verantwortlich sind und dieser deshalb erst nach einem Intervall mit Schlaf auftritt.

Insgesamt nahmen 80 gesunde männliche Studenten im Alter von 18-30 Jahren als Probanden an der Schlafstudie teil. Jeweils die Hälfte der Probanden gehörte der Test- bzw. der Lernbedingung an. Alle Probanden kamen jeweils drei Nachmittage für Versuche in unser Labor. Zwei der Versuchstage waren parallel gestaltet und unterschieden sich lediglich durch ein 90-minütiges Wach- bzw. Nap- Intervall im Verlauf der Testung. Zu Beginn des Testtags wurden 36 Wortpaare (Deutsch – Fantasiesprache) präsentiert. In der folgenden Übungsphase wurde den Probanden der Testbedingung in drei Zyklen jeweils das deutsche Wort gezeigt und sie mussten das Wortpaar per Tastatureingabe korrekt vervollständigen. Geling dies nicht, wurde das ganze Wortpaar noch einmal präsentiert. Den Probanden der Lernbedingung wurden in der Übungsphase alle Wortpaare noch viermal präsentiert. Im Anschluss an die Übungsphase erfolgte für alle Probanden ein Test (Baseline) aller Wortpaare ohne Feedback, der im weiteren Verlauf als Ausgangswert diente. Nach einem 90-minütigen Intervall mit Nap oder einem Film wurde erneut getestet (Test) und noch ein weiteres Mal nach vier Tagen (Retest) per webbasierter Abfrage. Anhand dieser drei Testwerte wurden die Vergessensraten über die Zeit ermittelt. Der dritte Termin im Labor diente als Kontrolle, die Probanden hielten einen 90-minütigen Nap ohne weitere Aufgaben. Während des Schlafs erfolgte jeweils eine polysomnographische Aufzeichnung und die Schlafstadien sowie Spindeln wurden erfasst und ausgewertet. Die statistische Analyse der Einflüsse von Schlaf und

wiederholtem Testen erfolgte mittels ANOVA mit Messwiederholungen mit den Innersubjektfaktoren Zeit und Nap sowie dem Zwischensubjektfaktor Bedingung (Testing vs. Lernen).

Die ANOVA mit Messwiederholungen und nachfolgende Tests der Innersubjektivitätskontraste ergaben für die Faktoren Zeit, Zeit x Nap (ausschließlich Test vs. Baseline), sowie Zeit x Bedingung (ausschließlich Retest vs. Baseline) signifikante Ergebnisse. Diesem statistischen Befund entspricht ein signifikanter Unterschied der Vergessensrate zwischen Probanden der Nap-Bedingung und Probanden der Wachbedingung am Testtag selbst, nicht jedoch nach 4 Tagen. Probanden der Testbedingung unterschieden sich hingegen nicht am Testtag, sondern erst beim Retest nach 4 Tagen signifikant von denen der Lernbedingung. Es fand sich keine Interaktion von Zeit x Nap x Bedingung. Tiefschlaf korrelierte positiv mit der relativen Erinnerungsleistung am Testtag. Die Spindelparameter zeigten hingegen keinerlei signifikante Korrelationen mit kognitiver Leistung.

Die vorbeschriebenen positiven Effekte von Schlaf und Testen konnten in der vorliegenden Studie repliziert werden. Der Schlafeffekt war direkt am Testtag sichtbar, verschwand aber im Verlauf der folgenden Tage. Der Testing Effekt trat wie erwartet nicht am Testtag, sondern erst mit einer Verzögerung bei dem erneuten Test nach vier Tagen auf. Eine Interaktion von Schlaf und Testing im Sinne der Hypothese konnte hingegen nicht festgestellt werden. Eine weitere Arbeitsgruppe fand bei einem 12h – Intervall mit Nachtschlaf oder Wachphase eine negative Interaktion i.S. von vermindertem Testing Effekt bei Schlaf (Bäumel et al. 2014). Gemeinsam betrachtet sind diese Ergebnisse mit der Bifurkationshypothese (Kornell et al. 2011) vereinbar. Diese besagt, dass im Rahmen des Testing nur korrekt abgerufene Items selektiv und deutlich verstärkt werden, während alle durch wiederholte Präsentation gelernte Items relativ schwächer verstärkt werden. Die Stärke der Erinnerung fällt für alle Items ungeachtet des Lernmodus mit einer konstanten Geschwindigkeit, wobei aufgrund der initial höheren Verstärkung nach einem längeren Intervall mehr getestete Items oberhalb der Abrufbarkeitsschwelle verbleiben. Es wäre von Interesse ein genaueres Bild über den zeitlichen Verlauf des Testing Effekts zu erhalten, indem man innerhalb einer einzelnen Studie die Testzeitpunkte über einen größeren Zeitraum moduliert, möglicherweise in Kombination mit längeren Wachphasen durch Schlafentzug.

Literaturverzeichnis

- Abott, E. E. (1909): On the analysis of the factor of recall in the learning process. In: *The Psychological Review: Monograph Supplements* 11 (1), S. 159–177.
- Ackermann, S.; Hartmann, F.; Papassotiropoulos, A.; de Quervain, Dominique JF; Rasch, B. (2015): No Associations between Interindividual Differences in Sleep Parameters and Episodic Memory Consolidation. In: *Sleep* 38 (6), S. 951–959.
- Addison, P. S. (2002): *The Illustrated Wavelet Transform Handbook*. Introductory theory and applications in science, engineering, medicine and finance. Bristol: IoP, Institut of Physics Publ.
- Arkes, H. R.; Joyner, C. A.; Pezzo, M. V.; Nash, J. G.; Siegel-Jacobs, K.; Stone, E. (1994): The Psychology of Windfall Gains. In: *Organizational Behavior and Human Decision Processes* 59 (3), S. 331–347.
- Aschoff, J. (1965): Circadian Rhythms in Man. In: *Science* 148 (3676), S. 1427–1432.
- Astori, S.; Wimmer, R. D.; Lüthi, A. (2013): Manipulating sleep spindles--expanding views on sleep, memory, and disease. In: *Trends Neurosci.* 36 (12), S. 738–748.
- Bäumli, K.-H. T.; Holterman, C.; Abel, M. (2014): Sleep can reduce the testing effect: It enhances recall of restudied items but can leave recall of retrieved items unaffected. In: *J Exp Psychol Learn Mem Cogn* 40 (6), S. 1568–1581.
- Bergmann, T. O.; Molle, M.; Diedrichs, J.; Born, J.; Siebner, H. R. (2012): Sleep spindle-related reactivation of category-specific cortical regions after learning face-scene associations. In: *Neuroimage* 59 (3), S. 2733–2742.
- Bjork, R. A.; Bjork, E. L. (1992): A new theory of disuse and an old theory of stimulus fluctuation. In: A. F. Healy, S. M. Kosslyn und R. M. Shiffrin (Hg.): *From learning processes to cognitive processes. Essays in Honor of William K. Estes. Volume 2*. Hillsdale: Lawrence Erlbaum, S. 35–67.
- Bodizs, R.; Kis, T.; Lazar, A. S.; Havran, L.; Rigo, P.; Clemens, Z.; Halasz, P. (2005): Prediction of general mental ability based on neural oscillation measures of sleep. In: *J Sleep Res* 14 (3), S. 285–292.
- Bonnet, M. H.; Gomez, S.; Wirth, O.; Arand, D. L. (1995): The use of caffeine versus prophylactic naps in sustained performance. In: *Sleep* 18 (2), S. 97–104.
- Borbély, A. A.; Achermann, P. (1992): Concepts and models of sleep regulation: an overview. In: *Journal of Sleep Research* 1 (2), S. 63–79.
- Born, J.; Rasch, B.; Gais, S. (2006): Sleep to remember. In: *Neuroscientist* 12 (5), S. 410–424.
- Brooks, A.; Lack, L. (2006): A brief afternoon nap following nocturnal sleep restriction: which nap duration is most recuperative? In: *Sleep* 29 (6), S. 831–840.
- Butler, A. C.; Roediger, H. L. (2007): Testing improves long-term retention in a simulated classroom setting. In: *European Journal of Cognitive Psychology* 19 (4-5), S. 514–527.
- Butler, A. C.; Roediger, H. L. 3. (2008): Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. In: *Mem Cognit* 36 (3), S. 604–616.

- Buysse, D. J.; Reynolds, C. F. 3.; Monk, T. H.; Berman, S. R.; Kupfer, D. J. (1989): The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. In: *Psychiatry research* 28 (2), S. 193–213.
- Chan, J. C. K.; LaPaglia, J. A. (2011): The dark side of testing memory: Repeated retrieval can enhance eyewitness suggestibility. In: *Journal of Experimental Psychology: Applied* 17 (4), S. 418–432.
- Chan, J. C. K.; McDermott, K. B.; Roediger, H. L. (2006): Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. In: *Journal of Experimental Psychology: General* 135 (4), S. 553–571.
- Congleton, A.; Rajaram, S. (2012): The origin of the interaction between learning method and delay in the testing effect: The roles of processing and conceptual retrieval organization. In: *Mem Cogn* 40 (4), S. 528–539.
- Cox, R.; Hofman, W. F.; Talamini, L. M. (2012): Involvement of spindles in memory consolidation is slow wave sleep-specific. In: *Learning & Memory* 19 (7), S. 264–267.
- Crowley, K. (2011): Sleep and sleep disorders in older adults. In: *Neuropsychology review* 21 (1), S. 41–53.
- Dang-Vu, T. T.; McKinney, S. M.; Buxton, O. M.; Solet, J. M.; Ellenbogen, J. M. (2010): Spontaneous brain rhythms predict sleep stability in the face of noise. In: *Current Biology* 20 (15), R626-7.
- Diekelmann, S.; Born, J. (2010): The memory function of sleep. In: *Nat Rev Neurosci* 11 (2), S. 114–126.
- Diekelmann, S.; Wilhelm, I.; Born, J. (2009): The whats and whens of sleep-dependent memory consolidation. In: *Sleep Med Rev* 13 (5), S. 309–321.
- Dinges, D. F.; Powell, J. W. (1985): Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. In: *Behavior Research Methods, Instruments, & Computers* 17 (6), S. 652–655.
- Dresler, M. (Hg.) (2007): *Wissenschaft an den Grenzen des Verstandes. Beiträge aus den Natur- und Lebenswissenschaften; eine Publikation des MinD-Hochschul-Netzwerkes.* Stuttgart: Hirzel.
- Dresler, M.; Spoormaker, V. I.; Beiting, P.; Czisch, M.; Kimura, M.; Steiger, A.; Holsboer, F. (2014): Neuroscience-driven discovery and development of sleep therapeutics. In: *Pharmacol. Ther.* 141 (3), S. 300–334.
- Dudai, Y. (2004): The Neurobiology of Consolidations, Or, How Stable is the Engram? In: *Annu. Rev. Psychol.* 55 (1), S. 51–86.
- 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. In: *Psychological Science in the Public Interest* 14 (1), S. 4–58.
- Ebbinghaus, H. (1885): *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie.* Duncker & Humblot.

- Ellenbogen, J. M.; Hulbert, J. C.; Stickgold, R.; Dinges, D. F.; Thompson-Schill, S. L. (2006a): Interfering with Theories of Sleep and Memory: Sleep, Declarative Memory, and Associative Interference. In: *Current Biology* 16 (13), S. 1290–1294.
- Ellenbogen, J. M.; Payne, J. D.; Stickgold, R. (2006b): The role of sleep in declarative memory consolidation: passive, permissive, active or none? In: *Curr Opin Neurobiol* 16 (6), S. 716–722.
- Ficca, G.; Axelsson, J.; Mollicone, D. J.; Muto, V.; Vitiello, M. V. (2010): Naps, cognition and performance. In: *Sleep Medicine Reviews* 14 (4), S. 249–258.
- Ficca, G.; Lombardo, P.; Rossi, L.; Salzarulo, P. (2000): Morning recall of verbal material depends on prior sleep organization. In: *Behavioural Brain Research* 112 (1-2), S. 159–163.
- Fogel, S. M.; Nader, R.; Cote, K. A.; Smith, C. T. (2007): Sleep spindles and learning potential. In: *Behavioral Neuroscience* 121 (1), S. 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. In: *Neurosci Biobehav Rev* 35 (5), S. 1154–1165.
- Fowler, M. J.; Sullivan, M. J.; Ekstrand, B. R. (1973): Sleep and memory. In: *Science* 179 (4070), S. 302–304.
- Gaillard, J. M.; Blois, R. (1981): Spindle density in sleep of normal subjects. In: *Sleep* 4 (4), S. 385–391.
- Gais, S. (2006): Sleep after learning aids memory recall. In: *Learning & Memory* 13 (3), S. 259–262.
- Gais, S.; Molle, M.; Helms, K.; Born, J. (2002): Learning-dependent increases in sleep spindle density. In: *J Neurosci* 22 (15), S. 6830–6834.
- Galland, B. C.; Taylor, B. J.; Elder, D. E.; Herbison, P. (2012): Normal sleep patterns in infants and children: a systematic review of observational studies. In: *Sleep Medicine Reviews* 16 (3), S. 213–222.
- Gennaro, L. de; Ferrara, M. (2003): Sleep spindles: an overview. In: *Sleep Med Rev* 7 (5), S. 423–440.
- Genzel, L.; Dresler, M.; Cornu, M.; Jager, E.; Konrad, B.; Adamczyk, M. et al. (2015a): Medial prefrontal-hippocampal connectivity and motor memory consolidation in depression and schizophrenia. In: *Biological Psychiatry* 77 (2), S. 177–186.
- Genzel, L.; Kiefer, T.; Renner, L.; Wehrle, R.; Kluge, M.; Grözinger, M. et al. (2012): Sex and modulatory menstrual cycle effects on sleep related memory consolidation. In: *Psychoneuroendocrinology* 37 (7), S. 987–998.
- Genzel, L.; Kroes, Marijn C W; Dresler, M.; Battaglia, F. P. (2014): Light sleep versus slow wave sleep in memory consolidation: a question of global versus local processes? In: *Trends Neurosci.* 37 (1), S. 10–19.
- Genzel, L.; Spoormaker, V. I.; Konrad, B. N.; Dresler, M. (2015b): The role of rapid eye movement sleep for amygdala-related memory processing. In: *Neurobiology of Learning and Memory* 122, S. 110–121.
- Glover, J. A. (1989): The "testing" phenomenon: Not gone but nearly forgotten. In: *Journal of Educational Psychology* 81 (3), S. 392–399.

- Greenhouse, S. W.; Geisser, S. (1959): On methods in the analysis of profile data. In: *Psychometrika* 24 (2), S. 95–112.
- Hoddes, E.; Zarcone, V.; Smythe, H.; Phillips, R.; Dement, W. C. (1973): Quantification of Sleepiness: A New Approach. In: *Psychophysiology* 10 (4), S. 431–436.
- Horne, J. A.; Ostberg, O. (1976): A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. In: *International journal of chronobiology* 4 (2), S. 97–110.
- Iber, C.; Ancoli-Israel, S.; Chesson, A.; Quan, S. F. (2007): Das AASM-Manual zum Scoring von Schlaf und assoziierten Ereignissen. Regeln, Terminologie und technische Spezifikationen. [Berlin]: Steinkopff.
- James, W. (1890): *The principles of psychology*, Vol 1. New York: Henry Holt und Co.
- Jenkins, J. G.; Dallenbach, K. M. (1924): Obliviscence during Sleep and Waking. In: *The American Journal of Psychology* 35 (4), S. 605–612.
- Karpicke, J. D.; Blunt, J. R. (2011): Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. In: *Science* 331 (6018), S. 772–775.
- Karpicke, J. D.; Butler, A. C.; Roediger III, H. L. (2009): Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? In: *Memory* 17 (4), S. 471–479.
- Karpicke, J. D.; Roediger, H. L. (2008): The Critical Importance of Retrieval for Learning. In: *Science* 319 (5865), S. 966–968.
- Killgore, W. D. S. (2010): Effects of sleep deprivation on cognition. In: *Prog. Brain Res.* 185, S. 105–129.
- Kornell, N.; Bjork, R. A.; Garcia, M. A. (2011): Why tests appear to prevent forgetting: A distribution-based bifurcation model. In: *Journal of Memory and Language* 65 (2), S. 85–97.
- Kromann, C. B.; Jensen, M. L.; Ringsted, C. (2009): The effect of testing on skills learning. In: *Medical Education* 43 (1), S. 21–27.
- Krosigk, M. von; Bal, T.; McCormick, D. (1993): Cellular mechanisms of a synchronized oscillation in the thalamus. In: *Science* 261 (5119), S. 361–364.
- Kumar, V. K.; Kemmler, D.; Holman, E. R. (1997): The Creativity Styles Questionnaire--Revised. In: *Creativity Research Journal* 10 (1), S. 51–58.
- Lahl, O.; Wispel, C.; Willigens, B.; Pietrowsky, R. (2008): An ultra short episode of sleep is sufficient to promote declarative memory performance. In: *J Sleep Res* 17 (1), S. 3–10.
- Lansink, C. S.; Goltstein, P. M.; Lankelma, J. V.; McNaughton, B. L.; Pennartz, Cyriel M A (2009): Hippocampus leads ventral striatum in replay of place-reward information. In: *PLoS biology* 7 (8), e1000173.
- Larsen, D. P.; Butler, A. C.; Roediger III, Henry L (2008): Test-enhanced learning in medical education. In: *Medical Education* 42 (10), S. 959–966.
- Lechner, H. A.; Squire, L. R.; Byrne, J. H. (1999): 100 years of consolidation--remembering Müller and Pilzecker. In: *Learning & Memory* 6 (2), S. 77–87.
- Lehrl, S. (2005): *Manual zum MWT-B*. 5. Aufl. Balingen: Spitta-Verl.

- Loh, S.; Lamond, N.; Dorrian, J.; Roach, G.; Dawson, D. (2004): The validity of psychomotor vigilance tasks of less than 10-minute duration. In: *Behavior Research Methods, Instruments, & Computers* 36 (2), S. 339–346.
- Luppi, P.-H.; Clement, O.; Sapin, E.; Peyron, C.; Gervasoni, D.; Léger, L.; Fort, P. (2012): Brainstem mechanisms of paradoxical (REM) sleep generation. In: *Pflugers Arch.* 463 (1), S. 43–52.
- Marisch, C. C. (2015): Auswirkungen von Schlaf und dessen polysomnographischen Korrelaten auf kreative Prozesse. Dissertation, LMU München: Medizinische Fakultät.
- McDaniel, M. A.; Roediger, H. L. 3.; McDermott, K. B. (2007): Generalizing test-enhanced learning from the laboratory to the classroom. In: *Psychon Bull Rev* 14 (2), S. 200–206.
- McDermott, K. B. (2006): Paradoxical effects of testing: repeated retrieval attempts enhance the likelihood of later accurate and false recall. In: *Mem Cognit* 34 (2), S. 261–267.
- McGaugh, J. L. (1966): Time-Dependent Processes in Memory Storage. In: *Science* 153 (3742), S. 1351–1358.
- Mednick, S. C.; McDevitt, E. A.; Walsh, J. K.; Wamsley, E.; Paulus, M.; Kanady, J. C.; Drummond, Sean P A (2013): The critical role of sleep spindles in hippocampal-dependent memory: a pharmacology study. In: *J Neurosci* 33 (10), S. 4494–4504.
- Medrick, S. C.; Drummond, S. P. A. (2009): Napping. In: L. R. Squire (Hg.): *Encyclopedia of Neuroscience*. Oxford: Academic Press, S. 1–6. Online verfügbar unter <http://www.sciencedirect.com/science/article/pii/B9780080450469000759>.
- Milner, C. E.; Cote, K. A. (2009): Benefits of napping in healthy adults: impact of nap length, time of day, age, and experience with napping. In: *J Sleep Res* 18 (2), S. 272–281.
- Morris, C. D.; Bransford, J. D.; Franks, J. J. (1977): Levels of processing versus transfer appropriate processing. In: *Journal of verbal learning and verbal behavior* 16 (5), S. 519–533.
- Moser, D.; Anderer, P.; Gruber, G.; Parapatics, S.; Loretz, E.; Boeck, M. et al. (2009): Sleep classification according to AASM and Rechtschaffen & Kales: effects on sleep scoring parameters. In: *SLEEP* 32 (2), S. 139–149.
- Mulrine, H. M.; Signal, T. L.; Berg, Margo J. van den; Gander, P. H. (2012): Post-Sleep Inertia Performance Benefits of Longer Naps in Simulated Nightwork and Extended Operations. In: *Chronobiol Int* 29 (9), S. 1249–1257.
- Oldfield, R. C. (1971): The assessment and analysis of handedness: The Edinburgh inventory. In: *Neuropsychologia* 9 (1), S. 97–113.
- Payne, J. D.; Stickgold, R.; Swanberg, K.; Kensinger, E. A. (2008): Sleep preferentially enhances memory for emotional components of scenes. In: *Psychol Sci* 19 (8), S. 781–788.
- Payne, J. D.; Tucker, M. A.; Ellenbogen, J. M.; Wamsley, E. J.; Walker, M. P.; Schacter, D. L. et al. (2012): Memory for Semantically Related and Unrelated Declarative Information: The Benefit of Sleep, the Cost of Wake. In: *PLoS ONE* 7 (3), e33079.
- Piosczyk, H.; Kloepfer, C.; Riemann, D.; Nissen, C. (2009): Schlaf, Plastizität und Gedächtnis. In: *Somnologie* 13 (1), S. 43–51.
- Plihal, W.; Born, J. (1997): Effects of early and late nocturnal sleep on declarative and procedural memory. In: *J Cogn Neurosci* 9 (4), S. 534–547.

Pyc, M. A.; Rawson, K. A. (2009): Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? In: *Journal of Memory and Language* 60 (4), S. 437–447.

Pyc, M. A.; Rawson, K. A. (2010): Why Testing Improves Memory: Mediator Effectiveness Hypothesis. In: *Science* 330 (6002), S. 335.

Qin, Y.-L.; McNaughton, B. L.; Skaggs, W. E.; Barnes, C. A. (1997): Memory reprocessing in corticocortical and hippocampocortical neuronal ensembles. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 352 (1360), S. 1525–1533.

Quintilian, M. (1995): Ausbildung des Redners; zwölf Bücher. Hrsg. und übers. von Helmut Rahn. 2 Bände. Darmstadt: Wissenschaftliche Buchgesellschaft.

Rasch, B.; Born, J. (2013): About Sleep's Role in Memory. In: *Physiological Reviews* 93 (2), S. 681–766.

Rasch, B.; Buchel, C.; Gais, S.; Born, J. (2007): Odor cues during slow-wave sleep prompt declarative memory consolidation. In: *Science* 315 (5817), S. 1426–1429.

Rattenborg, N. C.; Lesku, J. A.; Martinez-Gonzalez, D.; Lima, S. L. (2007): The non-trivial functions of sleep. In: *Sleep Med Rev* 11 (5), S. 405–409.

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

Roediger, H. L.; Agarwal, P. K.; McDaniel, M. A.; McDermott, K. B. (2011): Test-enhanced learning in the classroom: Long-term improvements from quizzing. In: *Journal of Experimental Psychology: Applied* 17 (4), S. 382–395.

Roediger, H. L.; Butler, A. C. (2011): The critical role of retrieval practice in long-term retention. In: *Trends in Cognitive Sciences* 15 (1), S. 20–27.

Roediger, H. L.; Karpicke, J. D. (2006a): Test-enhanced learning: taking memory tests improves long-term retention. In: *Psychol Sci* 17 (3), S. 249–255.

Roediger, H. L.; Karpicke, J. D. (2006b): The Power of Testing Memory. In: *Perspect on Psych Science* 1 (3), S. 181–210.

Roenneberg, T.; Kumar, C. J.; Mewes, M. (2007): The human circadian clock entrains to sun time. In: *Curr Biol* 17 (2), R44-5.

Saper, C. B.; Scammell, T. E.; Lu, J. (2005): Hypothalamic regulation of sleep and circadian rhythms. In: *Nature* 437 (7063), S. 1257–1263.

Schabus, M.; Gruber, G.; Parapatics, S.; Sauter, C.; Klosch, G.; Anderer, P. et al. (2004): Sleep spindles and their significance for declarative memory consolidation. In: *Sleep* 27 (8), S. 1479–1485.

Schabus, M.; Hodlmoser, K.; Pecherstorfer, T.; Klosch, G. (2005): Influence of Midday Naps on Declarative Memory Performance and Motivation. Der Einfluss von Mittagsschläpfchen auf deklarative Gedächtnisleistung und Motivation. In: *Somnologie* 9 (3), S. 148–153.

Schacter, D. L. (2001): The seven sins of memory. How the mind forgets and remembers. Boston: Houghton Mifflin.

- Schmidt, C. (2006): Encoding Difficulty Promotes Postlearning Changes in Sleep Spindle Activity during Napping. In: *Journal of Neuroscience* 26 (35), S. 8976–8982.
- Schmidt, R. F. (Hg.) (2007): Physiologie des Menschen. Mit Pathophysiologie. 30., neu bearb. und aktualisierte Aufl. Heidelberg: Springer (Springer-Lehrbuch).
- Schmitt, M.; Altstötter-Gleich, C.; Hinz, A.; Maes, J.; Brähler, E. (2006): Normwerte für das Vereinfachte Beck-Depressions-Inventar (BDI-V) in der Allgemeinbevölkerung. In: *Diagnostica* 52 (2), S. 51–59.
- Schmitt, M.; Maes, J. (2000): Vorschlag zur Vereinfachung des Beck-Depressions-Inventars (BDI). In: *Diagnostica* 46 (1), S. 38–46.
- Squire, L. R.; Zola, S. M. (1996): Structure and function of declarative and nondeclarative memory systems. In: *Proc Natl Acad Sci U S A* 93 (24), S. 13515–13522.
- Steriade, M. (2000): Corticothalamic resonance, states of vigilance and mentation. In: *Neuroscience* 101 (2), S. 243–276.
- Steriade, M.; Deschenes, M.; Domich, L.; Mulle, C. (1985): Abolition of spindle oscillations in thalamic neurons disconnected from nucleus reticularis thalami. In: *Journal of neurophysiology* 54 (6), S. 1473–1497.
- Stickgold, R.; Walker, M. P. (2013): Sleep-dependent memory triage: evolving generalization through selective processing. In: *Nat Neurosci* 16 (2), S. 139–145.
- Stuck, B. A.; Fricke-Oerkermann, L. (2013): Praxis der Schlafmedizin. Schlafstörungen bei Erwachsenen und Kindern. Diagnostik, Differentialdiagnostik und Therapie. 2. Aufl. Berlin: Springer.
- Sumowski, J. F.; Leavitt, V. M.; Cohen, A.; Paxton, J.; Chiaravalloti, N. D.; DeLuca, J. (2013): Retrieval practice is a robust memory aid for memory-impaired patients with MS. In: *Multiple Sclerosis Journal* 19 (14), S. 1943–1946.
- Sutherland, G. R.; McNaughton, B. (2000): Memory trace reactivation in hippocampal and neocortical neuronal ensembles. In: *Current Opinion in Neurobiology* 10 (2), S. 180–186.
- Szpunar, K. K.; McDermott, K. B.; Roediger, H. L. (2008): Testing during study insulates against the buildup of proactive interference. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34 (6), S. 1392–1399.
- Takashima, A.; Petersson, K. M.; Rutters, F.; Tendolkar, I.; Jensen, O.; Zwarts, M. J. et al. (2006): Declarative memory consolidation in humans: a prospective functional magnetic resonance imaging study. In: *Proc Natl Acad Sci U S A* 103 (3), S. 756–761.
- Tononi, G.; Cirelli, C. (2006): Sleep function and synaptic homeostasis. In: *Sleep Medicine Reviews* 10 (1), S. 49–62.
- Toppino, T. C.; Cohen, M. S. (2009): The testing effect and the retention interval: questions and answers. In: *Experimental psychology* 56 (4), S. 252–257.
- Tucker, M.; Hirota, Y.; Wamsley, E.; Lau, H.; Chaklader, A.; Fishbein, W. (2006): A daytime nap containing solely non-REM sleep enhances declarative but not procedural memory. In: *Neurobiology of Learning and Memory* 86 (2), S. 241–247.

- Tucker, M. A.; Fishbein, W. (2008): Enhancement of declarative memory performance following a daytime nap is contingent on strength of initial task acquisition. In: *Sleep* 31 (2), S. 197–203.
- Ujma, P. P.; Konrad, B. N.; Genzel, L.; Bleifuss, A.; Simor, P.; Potari, A. et al. (2014): Sleep Spindles and Intelligence: Evidence for a Sexual Dimorphism. In: *Journal of Neuroscience* 34 (49), S. 16358–16368.
- van den Broek, G. S. E.; Segers, E.; Takashima, A.; Verhoeven, L. (2014): Do testing effects change over time? Insights from immediate and delayed retrieval speed. In: *Memory* 22 (7), S. 803–812.
- Vertes, R. P. (2004): Memory consolidation in sleep; dream or reality. In: *Neuron* 44 (1), S. 135–148.
- Vertes, R. P.; Siegel, J. M. (2005): Time for the sleep community to take a critical look at the purported role of sleep in memory processing. In: *Sleep* 28 (10), S. 1228–1229.
- Wagner, U.; Gais, S.; Haider, H.; Verleger, R.; Born, J. (2004): Sleep inspires insight. In: *Nature* 427 (6972), S. 352–355.
- Walker, M. P.; Stickgold, R. (2006): Sleep, Memory, and Plasticity. In: *Annu. Rev. Psychol.* 57 (1), S. 139–166.
- Wauquier, A. (1993): Aging and changes in phasic events during sleep. In: *Physiology & behavior* 54 (4), S. 803–806.
- Weiss, R.; Weiss, B. (2006): CFT-20R Grundintelligenztest Skala, Revision 2. Göttingen, The Netherlands: Hogrefe Verlag.
- Wheeler, M.; Ewers, M.; Buonomano, J. (2003): Different rates of forgetting following study versus test trials. In: *Memory* 11 (6), S. 571–580.
- Wilhelm, I.; Diekelmann, S.; Molzow, I.; Ayoub, A.; Mölle, M.; Born, J. (2011): Sleep Selectively Enhances Memory Expected to Be of Future Relevance. In: *The Journal of Neuroscience* 31 (5), S. 1563–1569.
- Wilson, M. A.; McNaughton, B. L. (1994): Reactivation of hippocampal ensemble memories during sleep. In: *Science* 265 (5172), S. 676–679.
- Wixted, J. T. (2004): The psychology and neuroscience of forgetting. In: *Annu Rev Psychol* 55, S. 235–269.
- Wyatt, J. K.; Ritz-De Cecco, A.; Czeisler, C. A.; Dijk, D. J. (1999): Circadian temperature and melatonin rhythms, sleep, and neurobehavioral function in humans living on a 20-h day. In: *Am J Physiol* 277 (4 Pt 2), R1152-63.
- Yaroush, R.; Sullivan, M. J.; Ekstrand, B. R. (1971): Effect of sleep on memory. II. Differential effect of the first and second half of the night. In: *J Exp Psychol* 88 (3), S. 361–366.
- Zaromb, F. M.; Roediger, H. L. (2010): The testing effect in free recall is associated with enhanced organizational processes. In: *Mem Cogn* 38 (8), S. 995–1008.
- Zimbardo, P. G.; Gerrig, R. J. (2008): Psychologie. 18., aktual. Aufl. München: Pearson Studium (Pearson Studium - Psychologie).

Anhang

In der Lernaufgabe verwendete Wortlisten:

Wortliste 1		Wortliste 2		
Reiter	rustar	1	Schule	scheb
Feuer	apit	2	Turm	ripat
Tier	serise	3	Meer	manel
Strasse	sarre	4	Zeitung	krustele
Weber	krospe	5	Riese	sirke
Heer	sahe	6	Mutter	uma
Möbel	knarsch	7	Insekt	ramete
Freund	patrum	8	Fluss	tinxur
Vogel	storl	9	Hochschule	pame
Feier	gelerge	10	Schmied	ranscher
Reptil	kreik	11	Wohnung	interan
Regen	enser	12	Infektion	nesa
Gelenk	floschen	13	Getreide	kulinse
Garten	surke	14	Allee	nelle
Bündnis	bonu	15	Diamant	sturk
Fabrik	ampe	16	Sonate	alluse
Pflanze	kanabus	17	Buch	redor
Fahne	kirse	18	Katze	bisch
Zügel	schern	19	Angriff	amaki
Werbung	blisch	20	Puppe	beni
Pächter	salzal	21	Eisenbahn	terion
Berg	padir	22	Küche	tomandel
Herrscher	ischab	23	Landschaft	tital
Schauspiel	dramin	24	Musiker	rocktur
Gebäude	konturas	25	Kleidung	fraktan
Kirche	kelume	26	Familie	katesa
Krankheit	kerete	27	Auto	irtax
Reise	kukutur	28	Gefängnis	schrunk
Revolver	kenekel	29	Orkan	rotation
Unterwelt	famul	30	Küste	kulerane
Blasinstrument	eleume	31	Blumenstrauß	strun
Gletscher	krusse	32	Flasche	rezirne
Glaube	orakium	33	Gruppe	gesonk
Mädchen	bien	34	Körper	beinil
Tanne	tronke	35	Hafen	nale
Zirkus	struke	36	Maler	akifer

Danksagung

Ich danke an erster Stelle Prof. Dr. Dr. Dr. h.c. mult. Florian Holsboer, der es mir ermöglicht hat am Institut zu arbeiten.

Großer Dank gilt meinem Doktorvater Prof. Dr. Axel Steiger für die Überlassung des Themas, die Einführung in die Welt der Schlafforschung und seine Begleitung auf dem Weg zur Promotion.

Meinem Betreuer Dr. Martin Dresler möchte ich an dieser Stelle ganz besonders danken, er hat mich während der gesamten Zeit gut und zuverlässig betreut und stand mir bei Fragen jeder Art stets zügig und kompetent mit gutem Rat zur Seite. Darüber hinaus danke ich ihm für seine Unterstützung und Beratung bei der Durchführung und Interpretation der statistischen Analysen.

Mein Dank gilt auch Dr. Lisa Genzel, die mich auf das Thema aufmerksam gemacht hat und Dr. Boris Konrad für die Programmierung der Webseite und spannende Einblicke in die Welt der Mnemotechniken.

Bedanken möchte ich mich bei Gabriele Kohl, Luise Vogel, Birte Balzer, Beate Riemenschneider und Robert Neuner für die stets freundliche Unterstützung und kollegiale Zusammenarbeit. Ein Dankeschön geht auch an Annabell Bleifuss und Thomas Rost für ihre Unterstützung sowie an alle Mitarbeiter der AG Steiger für den freundlichen kollegialen Austausch.

Ein Dankeschön auch an die VolkswagenStiftung für die finanzielle Förderung des Forschungsprojekts.

Ein besonderer Dank gilt allen Probanden, die bei der Studie mitgemacht haben, und dafür auch am Wochenende schon morgens früh aufgestanden sind. Nur durch ihre zuverlässige Mitarbeit war die Studie möglich.

Ein großes Dankeschön auch an meine drei Mitstreiterinnen Cynthia Marisch, Johanna Pömmerl und Fee Stremmel. Die Arbeit im Team hat viel Spaß gemacht und wir haben uns bei der Durchführung der Studie wunderbar ergänzt.

Zu guter Letzt möchte ich meinen Eltern und meiner Frau danken, die mir in dieser Zeit unermüdlich zur Seite standen und mich stets unterstützt und ermuntert haben.

Eidesstattliche Versicherung

Stintzing, Johannes

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

Wirkung von Nachmittagsschlaf und wiederholtem Abruf auf die verbale Gedächtniskonsolidierung

selbstständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, den 29.07.2016

Ort, Datum

Unterschrift Doktorandin/Doktorand