Der Effekt von Schlaf auf das Gedächtnis

Inauguraldissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Olaf Lahl

geboren in Neuss

April 2006

Aus dem Institut für Experimentelle Psychologie

der Heinrich-Heine-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung

der Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Heine-Universität Düsseldorf

Referent:	Prof. Dr. Reinhard Pietrowsky
Koreferent:	Prof. Dr. Petra Stoerig

Tag der mündlichen Prüfung: 23.06.2006

Inhalt

1. Zusammenfassung

Die Frage „Warum schlafen wir?“ kann von der Wissenschaft bislang nicht befriedigend beantwortet werden. Eine der in diesem Zusammenhang viel diskutierten Hypothesen schlägt vor, dass Prozesse der Gedächtnisbildung im Schlaf optimiert werden. Typischerweise zeigen Probanden, die zwischen Erwerb und Abruf eines bestimmten Lernmaterials geschlafen haben, eine bessere Behaltensleistung für die gelernten Inhalte als Probanden, die während des Retentionsintervalls wach blieben. Bislang ist allerdings nicht geklärt, ob dieser verbesserten Gedächtnisleistung tatsächlich ein funktionell bedeutsamer Effekt zugrunde liegt, der eine weit reichende und robuste Optimierung unterschiedlicher Funktionsbereiche des Gedächtnisses vermittelt. In der vorliegenden Arbeit wurde daher die Wirkung von Schlaf auf das Gedächtnis unter verschiedenen Bedingungen getestet, um Aufschluss über die Reichweite des Effektes zu erhalten.

In zwei Experimenten wurde geprüft, inwieweit der beschriebene Effekt von (i) circadianen Faktoren der mit Schlaf einhergehenden Nachtzeit, (ii) von der Art des Abrufs aus dem Langzeitgedächtnis und (iii) der Zeitdauer zwischen Erwerb und Wiedergabe des Lernmaterials abhängt. Zwei weitere Arbeiten gingen der Frage nach, wie sich Schlaf auf (iv) irrelevante Informationen und (v) Trugerinnerungen auswirkt. Aus den Ergebnissen geht hervor, dass (i) die Nachtzeit allein, unabhängig von Schlaf, zu einer Verbesserung der Behaltensleistung führen kann, (ii) unterschiedliche Methoden des Abrufs keinen Einfluss haben, (iii) ein positiver Effekt auf die Behaltensleistung nicht unmittelbar, sondern erst mit zeitlicher Verzögerung einsetzen kann und auch bei der Speicherung (iv) irrelevanter und (v) falscher Erinnerungen wirksam sein kann. Im Hinblick auf die Ausgangsfragestellung nach funktionellen Aspekten des Schlaf-Gedächtnis-Effektes sprechen die (i) Auslösbarkeit durch Nachtzeit allein und die (ii) empfindliche Abhängigkeit von der Dauer des Behaltensintervalls sowie die (iv, v) mangelnde Selektivität gegenüber unerwünschten Inhalten gegen die Annahme spezifischer Prozesse, die funktionell mit Schlaf in Verbindung stehen.

In einem zusätzlichen Methodenbeitrag wurde das Instrumentarium der kognitiven Schlafforschung um ein Programmpaket erweitert, mit dessen Anwendung sich verbale Lernmaterialien für die Anwendung von Messwiederholungsplänen automatisiert generieren lassen, die semantisch unterschiedlich, aber psycholinguistisch äquivalent sind.

2. Einleitung

Anlässlich seines 125. Jubiläums formulierte das Science Magazine einen Katalog von 125 Fragestellungen, auf die die modernen Naturwissenschaften bislang keine Antwort gefunden haben. Neben den klassischen Problemen einer Vereinheitlichung der Physik oder der Möglichkeit extraterrestrischen Lebens stößt man in dieser Aufzählung alsbald auch auf die Frage: „Why do we sleep?". Auch wenn die allnächtlich wiederkehrende imperative Müdigkeit des gesunden Schläfers eine rasche Antwort im Sinne einer irgendwie gearteten Erholungsfunktion nahe legen mag, hat sich diese Idee bisher wissenschaftlich kaum bewähren können. Sie teilt dieses Schicksal mit zahlreichen weiteren Vorschlägen, denen es bislang ungeachtet ihrer inhaltlichen Attraktivität an einer ausreichenden empirischen Absicherung mangelt. Allen Kandidaten ist gemeinsam, dass sie bestimmte Teilaspekte der Schlafphänomenologie erklären können, andere hingegen nicht.

In dieser Arbeit soll ein Ansatz herausgegriffen werden, der in der Vergangenheit immerhin zahlreiche und teilweise auch sehr erfolgreiche empirische Bestätigungsversuche erfahren hat und damit als einer der viel versprechenden Nachfolger der derzeitigen Unklarheit gelten darf. Es ist dies die Idee, dass Schlaf eine zentrale Rolle für Lernen und Gedächtnis spielt. In dieser Konzeptualisierung besteht die oder zumindest eine Funktion des Schlafes in der Konsolidierung von zuvor flüchtigen Gedächtnisinhalten, die während des Wachzustandes aufgenommen wurden. Ihre theoretische Motivation liegt auf der Hand. Mit der Ausschaltung des Bewusstseins stellt Schlaf einen idealen Zustand für die „Offline" Verarbeitung neuer Inhalte ohne Interferenz durch gleichzeitig eintreffende konkurrierende Informationen dar.

In vier Studien hat der Verfasser versucht, Folgerungen, die sich aus dieser Überlegung ergeben, einer empirischen Testung zu unterziehen. In einer fünften Arbeit schließlich hat er das methodische Instrumentarium der kognitiven Schlafforschung um ein Werkzeug erweitert, mit dessen Hilfe bestimmte experimentelle Designs in diesem Bereich erst einwandfrei anwendbar werden. Die wesentlichen Ergebnisse dieser Arbeiten sollen im folgenden vorgestellt und dann im Überblick schlussfolgernd diskutiert werden.

3. Der „Sleep Effect“

Erste Hinweise auf einen günstigen Einfluss von Schlaf auf die Gedächtnisbildung gehen bereits aus den Arbeiten von Ebbinghaus (1885 / 1966) hervor. Als seine eigene Versuchsperson lernte er bekanntlich Listen sinnloser Silben bis zum Kriterium einer vollständigen Rezitation und maß nach Zeitintervallen von 20 Minuten bis 744 Stunden jeweils die Zeit, die er zum vollständigen Wiedererlernen der Listen benötigte. Die prozentuale Beschleunigung des Wiedererlernens im Vergleich zum ursprünglichen Lernen wertete Ebbinghaus als Behaltensindex und bezeichnete ihn als Ersparniswert. Es zeigte sich der mittlerweile klassische Verlauf einer zunächst sehr steil, dann immer langsamer abfallenden prozentualen Zeitersparnis s als Funktion der Zeit t (Abb. 1). Diese Verlaufsform lässt sich für $t > 0$ gut durch eine Potenzfunktion mit den Parametern α und λ annähern (Anderson & Schooler, 1991):

$$s(t) = a\, t^{-l},$$

wobei λ als Vergessensrate interpretiert werden kann, α ein Proportionalitätsfaktor ist und für $t = 0$ eine Ersparnis von $s(0) = 100\%$ definiert werden kann.

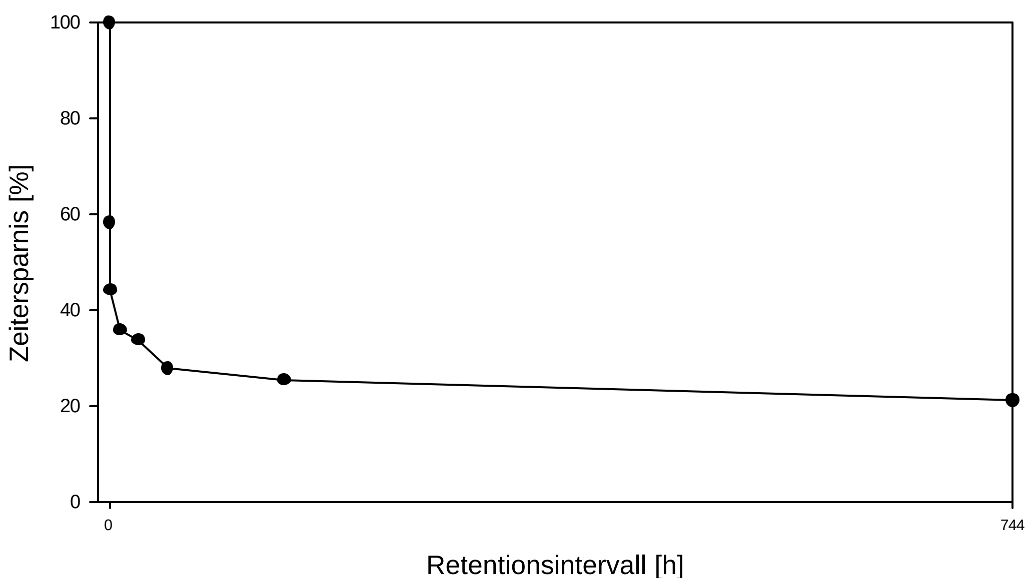


Abb. 1: Vergessenskurve nach Originaldaten von Ebbinghaus (1885/ 1966, S. 103).

Anschaulich zeigt sich die gute Übereinstimmung mit der angenommenen Potenzfunktion dann, wenn die beobachteten Daten im doppelt logarithmischen Koordinatensystem nur wenig von der Geraden mit dem Achsenabschnitt α und der Steigung –λ abweichen:

$$\log s(t) = \log a - l \log t$$

Mit der Methode der kleinsten Quadrate erhält man im doppelt logarithmischen Koordinatensystem die Parameterschätzungen $\hat{a}$ = 47.557 und $\hat{l}$ = 0.126. Wie Abb. 2 illustriert, gelingt die Anpassung an die entsprechende Regressionsgerade recht gut. Neben der ersten Messung weicht aber auch der Verlauf zwischen der dritten (8.8 h) und vierten Messung (24 h) nach oben ab.[1] Hier ist der Abfall der Zeitersparnis also geringer als von einer Potenzfunktion vorausgesagt. Von diesem 15-stündigen Zeitintervall berichtet Ebbinghaus aber, dass es seine nächtliche Schlafperiode beinhaltete. Ebbinghaus formulierte daraufhin zwar die „Annahme, dass Nacht und Schlaf […] die Abnahme der Nachwirkung [des Lernens, Anm. d. Verf.] ganz besonders verlangsamen“ (Ebbinghaus, 1885/1966, S. 104), ging dem Phänomen aber nicht weiter nach. Es handelt sich hierbei aber um einen ersten dokumentierten Hinweis darauf, dass Vergessensprozesse im Schlaf langsamer ablaufen könnten als im Wachzustand.

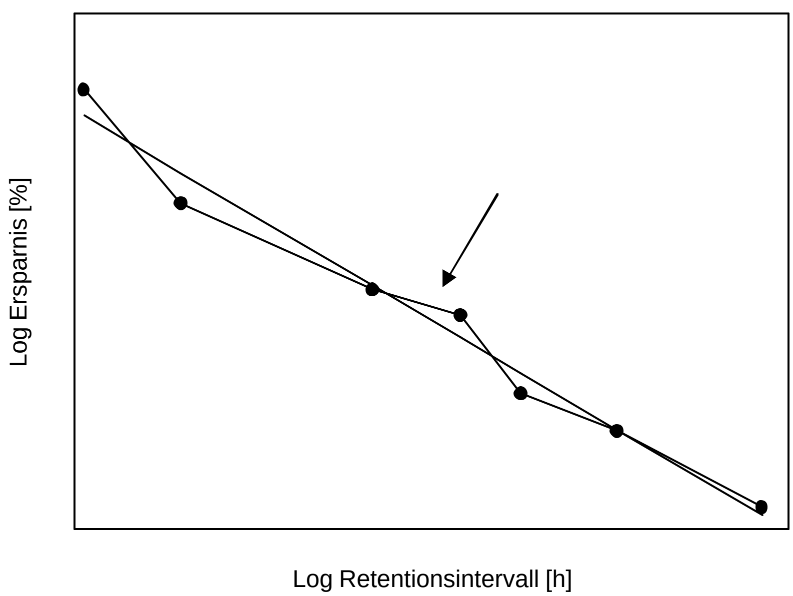


Abb. 2: Vergessenskurve aus Abb. 1 (Punkte) im doppelt logarithmischen Koordinatensystem. Im markierten Intervall zwischen der dritten und vierten Messung weicht der Verlauf der experimentellen Daten von der Regressionsgeraden log *s*(*t*) = log 47.557 – 0.126 log *t* nach oben ab.

[1] Da es sich um Abweichungen von einer Regressionsgeraden handelt, muss die Summe der Abweichungen natürlich Null ergeben. Bemerkenswert ist also nicht die Tatsache, dass es neben negativen auch positive Abweichungen gibt, sondern dass eine der beiden positiven Abweichungen auf die Schlafperiode fällt.

Es dauerte knapp vierzig Jahre, bis Jenkins und Dallenbach (1924) die beobachtete Anomalie explizit wieder aufgriffen, indem sie die Behaltensleistung für sinnlose Silben nach Retentionsintervallen von ein-, zwei-, vier- oder achtstündiger Dauer bestimmten, die ihre Versuchspersonen entweder mit wacher Tagesaktivität oder mit nächtlicher Schlafruhe verbracht hatten. Als Bestätigung der Ebbinghaus'schen Beobachtung ergab sich für alle vier Zeiträume eine bessere Behaltensleistung nach Schlaf, und diese Überlegenheit der Schlafbedingung nahm mit der Dauer des Retentionsintervalls sogar zu.

In der Folge gelang es mehreren Autoren, mittels unterschiedlicher Lernmaterialien (sinnlose Silben, einfache Wortlisten, Paarassoziationslisten, Kurzgeschichten) und Testmethoden (freier Abruf, gebundener Abruf, Neulernen, Rekognition) einen positiven Effekt von Schlaf auf die Behaltensleistung episodischer Inhalte nachzustellen (Benson & Feinberg, 1975, 1977; Dahl, 1928; Ekstrand, 1967, 1972; Ekstrand, Barrett, West & Maier, 1977; Lovatt & Warr, 1968; Nesca & Koulack, 1994; Newman, 1939; Spight, 1928; Van Ormer, 1932; Yaroush, Sullivan & Ekstrand, 1971), so dass sich Ekstrand et al. und später Cipolli (1995) sowie Ficca und Salzarulo (2004) veranlasst sahen, von *dem* „Sleep Effect" zu sprechen. Eine Erweiterung des Effektes auf nicht-episodische Inhalte konnte in der jüngeren Vergangenheit demonstriert werden. So ergaben sich auch für prozedurale Aufgaben höhere Behaltensindizes, wenn die Periode zwischen Lernen und erneuter Testung schlafend statt wach verbracht wurde. Gais, Plihal, Wagner und Born (2000) sowie Stickgold, Whidbee, Schirmer, Patel und Hobson (2000) konnten diesen Effekt für die visuelle Mustererkennung nachstellen. Fischer, Hallschmid, Elsner und Born (2002) sowie Walker, Brakefield, Morgan, Hobson und Stickgold (2002) gelang der Nachweis für die Nachahmung motorischer Sequenzen. Auch höhere kognitive Funktionen scheinen vom Schlaf zu profitieren. So wiesen Fenn, Nusbaum und Margoliash (2003) eine entsprechende schlafbedingte Erleichterung beim Erlernen einer synthetischen Sprache nach. Wagner, Gais, Haider, Verleger und Born (2004) konnten schließlich zeigen, dass Schlaf die Wahrscheinlichkeit erhöht, Einsicht in zuvor verborgene Regeln oder Schemata zu gewinnen.

4. Problemstellung

Trotz der zahlreichen konvergenten Einzelbefunde (weitere aktuelle Übersichten finden sich bei Blissit, 2001; Rauchs, Desgranges, Foret & Eustache, 2005; Stickgold, Hobson, R. Fosse & M. Fosse, 2001; Walker & Stickgold, 2004) wird die funktionelle Bedeutung des Schlafes für die Gedächtniskonsolidierung weiterhin kontrovers diskutiert (Horne, 2000; Stickgold & Walker, 2005; Vertes & Siegel, 2005a, 2005b). Neben einer Reihe diskrepanter tierexperimenteller Ergebnisse (Siegel, 2001; Vertes, 2004) ergeben sich Zweifel an der Robustheit und Allgemeinheit des Effektes aus weiteren Quellen.

So konnten Experimente, in denen circadiane Faktoren kontrolliert wurden, keinen Schlafeffekt auf die Gedächtnisleistung zeigen (Hockey, Davis & Gray, 1972; Nesca & Koulack, 1994). Diese Befunde werfen die Frage auf, ob Schlaf oder die circadiane Nachtphase, während derer der Schlaf üblicherweise stattfindet, die kritische Variable für die Gedächtniskonsolidierung bilden.

Ungeklärt ist auch, ob geringere Wiedergabeleistungen wacher Kontrollgruppen möglicherweise auf Abrufprobleme zurückgeführt werden müssen. Die Retentionsmessung beschränkt sich im allgemeinen auf einen einzelnen Abrufversuch, in dessen Verlauf das behaltene Material reproduziert werden soll. Untersuchungen mit wiederholten Abrufversuchen haben jedoch einen Anstieg der Erinnerungsleistung (Hypermnesie) demonstriert, wenn den Probanden in mehrfach hintereinander geschalteten Durchgängen Gelegenheit zur Wiedergabe gegeben wird (Lazar, 1969; Lazar & Van Laer, 1966; Payne, 1986; Richardson & Gropper, 1964). Offenbar unterschätzt die Beschränkung auf einen einzelnen Abrufversuch die Menge des behaltenen Materials, weil nicht alle Items direkt für den Zugriff zur Verfügung stehen. Auffällig sind in diesem Zusammenhang Ergebnisse der einzigen dem Verfasser bekannten Schlafstudie, in der wiederholte Tests zur Anwendung kamen. Barrett und Ekstrand (1972) fanden beim Vergleich zwischen Schlaf der ersten und zweiten Nachthälfte einen signifikant geringeren Verlust nach Schlaf der ersten Nachthälfte im ersten, nicht aber im zweiten Abrufversuch, da die Probanden der zweiten Nachthälfte im zweiten Test stärker aufgeholt hatten als ihre Kontrollen.

Von einer funktionellen Beteiligung des Schlafes an der Gedächtnisbildung würde man weiterhin eine langfristige Wirksamkeit erwarten. Die Effektivität des Behaltens und Erinnerns gegenüber Wachbedingungen sollte also nicht nur unmittelbar im Anschluss an die aktuelle Versuchsnacht gesteigert sein. Vielmehr sollte diese Überlegenheit im Verlauf der folgenden Tage und idealerweise auch darüber hinaus erhalten bleiben. Alternativ wäre ja auch denkbar, dass der Schlafeffekt nur eine kurzfristige, inzidentelle Wirkung darstellt, die mit fortschreitender Zeit wieder ausbleicht. Eine Entscheidung zwischen diesen beiden Ansätzen scheint auf der Grundlage der vorliegenden widersprüchlichen Befunde nicht möglich (Benson & Feinberg, 1975, 1977; Gibb, 1937, zitiert nach Richardson & Gough, 1963, S. 37; Graves, 1936; Idzikowski, 1984; Richardson & Gough, 1963): Von den insgesamt 17 untersuchten Retentionsintervallen mit einer Dauer zwischen 16 und 144 Stunden ergaben acht Intervalle eine Überlegenheit der Schlafbedingung, neun hingegen nicht.

Neben diesen methodischen Aspekten scheint der Generalisierbarkeit des Phänomens die Einseitigkeit seiner Untersuchung im Wege zu stehen. So hat die kognitive Schlafforschung zwar die exakte Reproduktion zuvor gelernter Information im Hinblick auf unterschiedlichste Materialien untersucht, sie hat sich damit aber auf den *reproduktiven* Aspekt der Gedächtnisleistung beschränkt. Demgegenüber scheinen ebenso zentrale Gedächtnisfunktionen wie die *Selektion* und *Konstruktion* von Informationen bislang keine Beachtung erfahren zu haben. So wurde die grundsätzliche Idee einer Funktion des Schlafes für das Filtern oder Löschen irrelevanter Informationen zwar mehrfach theoretisch formuliert (Crick & Mitchison, 1983; 1995; Evans & Newman, 1964; Feinberg & Evarts, 1969; Newman & Evans, 1965), aber nicht experimentell untersucht. Und während die Erforschung von rekonstruktiven Gedächtnisprozessen und Gedächtnistäuschungen in den letzten zehn Jahren einen massiven Auftrieb erfahren hat (vgl. Bruce & Winograd, 1998; Seamon et al., 2002), liegen bislang weder Konzepte noch Daten zur Wirkung von Schlaf auf Trugerinnerungen und mnestische Illusionen vor. Wenn Schlaf aber tatsächlich eine essentielle Bedeutung für die kognitive Anpassung des Organismus an seine Umwelt hat, dann ist zu erwarten, dass sich die Wirkung nicht auf die reproduktiven Gedächtnisleistungen beschränkt. Neben einer Verstärkung „erwünschter“ Inhalte sollte sich eine Abschwächung „unerwünschter“ Inhalte ergeben, die für die Interaktion mit der Umwelt keine Relevanz (mehr) haben oder denen eine Entsprechung in der Realität fehlt.

Eine entscheidende Bewährung im Popper'schen Sinne (Popper, 1968) könnte die Theorie einer funktionellen Bedeutung von Schlaf für das Gedächtnis also erfahren, wenn gezeigt werden könnte, dass es sich (i) nicht um ein Artefakt der Tageszeit oder eines mangelhaften Zugriffs handelt, wenn (ii) eine gewisse zeitliche Robustheit und Permanenz des Effektes demonstrierbar wäre und wenn schließlich (iii) eine funktionell adäquate Erweiterung auf gleichberechtigte Domänen des menschlichen Gedächtnisses gelänge. Mit den hier vorgestellten empirischen Arbeiten hat der Autor diesen Versuch unternommen. In der Arbeit „*Does the 'Sleep Effect' on Memory depend on Sleep or on Night Time?*“ (Lahl & Pietrowsky, eingereicht a) wird der Frage nachgegangen, ob die Behaltensleistung durch Schlaf oder durch Nachtzeit begünstigt wird und inwiefern Abruferleichterungen hierbei eine Rolle spielen. Die Arbeit „*The long range effect of sleep on episodic memory*” (Lahl & Pietrowsky, eingereicht b) widmet sich der zeitlichen Stabilität des Effektes. Die Studie „*Effect of sleep on task-irrelevant stimuli*“ (Lahl & Pietrowsky, eingereicht c) testet die Fragestellung, ob Schlaf das Vergessen irrelevanter Informationen begünstigt. Schließlich behandelt die Untersuchung „*Sleep enhances memory for things that never happened*“ (Lahl, Zlomuzica & Pietrowsky, eingereicht) den Effekt von Schlaf auf Erinnerungstäuschungen.

5. Methodischer Beitrag

Aufgrund ihrer leichten Handhabbarkeit und Quantifizierbarkeit gehören Listen von Worten oder Wortpaaren zu den verbreitetsten Materialien, wenn es um die Untersuchung des Gedächtnisses geht. Gleichzeitig ist die Verwendung von Messwiederholungsplänen, in denen jeder Proband nacheinander unterschiedliche experimentelle Bedingungen durchläuft, in zahlreichen Experimenten der kognitiven Schlafforschung anzutreffen (z. B. Drosopoulos, Wagner & Born, 2005; Ficca, Lombardo, Rossi & Salzarulo, 2000; Lewin & Glaubman, 1975; Meléndez et al., 2005; Plihal & Born, 1997, 1999a, 1999b). Um unerwünschte Lerneffekte von einer experimentellen Bedingung auf die folgende auszuschließen, bedient man sich dabei unterschiedlicher Worte je Bedingung. Da die Bedingungen andererseits nur hinsichtlich der unabhängigen Variable(n) voneinander abweichen sollen, müssen die unterschiedlichen Wortlisten in Bezug auf ihre Lern- und Erinnerbarkeit möglichst äquivalent sein. Damit ergibt sich die Notwendigkeit einer Parallelisierung der Wortlisten nach allen psycholinguistischen Attributen, die hierbei von Belang sein können. Eine nähere Sichtung der oben zitierten Experimente zeigt jedoch, dass hierbei stets nur ein Teil der wesentlichen Attribute berücksichtigt wurde. Ein Grund hierfür mag darin bestehen, dass sich die Parallelisierung bereits mit zwei zu berücksichtigenden Variablen als sehr schwierig gestaltet, wenn sie ohne Unterstützung durch Computer erfolgen soll. Da nach Kenntnis des Autors keine Software zur vollständigen und automatischen Parallelisierung von Wortlisten vorlag, hat er diese selbst entwickelt, öffentlich zugänglich gemacht und in der Publikation „*EQUIWORD: A software for the automatic creation of truly equivalent word lists*“ (Lahl & Pietrowsky, im Druck) dokumentiert.

6. Experiment I: Untersuchung des Schlaf-Gedächtnis-Effektes unter Kontrolle circadianer Faktoren

Der Vergleich zweier Bedingungen, von denen die eine den natürlichen Nachtschlaf (Schlaf/Nacht), die andere die gewohnte wache Tagesaktivität (Wach/Tag) der Probanden vorsieht, ist das übliche Vorgehen in Experimenten, in denen Schlaf die unabhängige Variable ist. Neben der gewünschten Manipulation des Faktors Schlaf wird mit diesem Design allerdings auch die unabhängige Variable Tageszeit variiert, so dass beide Faktoren letztlich konfundiert sind. Mit dem hier vorgestellten Experiment sollte die Fragestellung beantwortet werden, ob sich der positive Effekt von Schlaf auf die Behaltensleistung auch dann nachstellen lässt, wenn die angesprochene Konfundierung mit der Tageszeit vermieden wird.

Weiterhin ging es um die Frage, ob der Schlafeffekt auch dann erhalten bleibt, wenn Probleme beim Zugriff auf behaltenes Material minimiert werden. In diesem Zusammenhang wurde auf der Grundlage der verfügbaren Literatur (Lazar, 1969; Lazar & Van Laer, 1966) vermutet, dass wiederholte Abrufmöglichkeiten generell zu einem Zugewinn erinnerbarer Lerneinheiten führen würden, die in vorherigen Zugriffsversuchen zwar vorhanden, aber nicht abrufbar waren. Dieser Annahme liegt letztlich die Tulving'sche Unterscheidung zwischen *trace-dependent forgetting* (irreversibles Vergessen durch Spurenzerfall) und *cue-dependent forgetting* (vorübergehende Störung des Zugriffs auf prinzipiell vorhandene Inhalte aufgrund unzureichender Abrufhilfen) zugrunde (Tulving, 1974; Tulving & Madigan, 1970; Tulving & Psotka, 1971). Mögliche differenzielle Effekte der Schlaf/Wach-Bedingungen würden sich in unterschiedlichen Raten des Zugewinns im Verlauf der wiederholten Testtrials äußern.

Zur Beantwortung der ersten Fragestellung wurde den beiden herkömmlich eingesetzten Bedingungen Wach/Tag und Schlaf/Nacht eine zusätzliche Gruppe Wach/Nacht hinzugefügt. Zwischen Lernen und Wiedergabe war für jede Bedingung ein Retentionsintervall von sieben Stunden vorgesehen, das in der Wach/Tag-Gruppe mit gewohnter Tagesaktivität, in der Schlaf/Nacht-Gruppe mit Nachtschlaf und in der Wach/Nacht-Gruppe mit nächtlicher Wachheit (Schlafdeprivation) gefüllt war. Mit dem Vergleich zwischen den Gruppen Wach/Tag und Schlaf/Nacht sollte der klassische Schlafeffekt repliziert werden. Mit dem Gruppenvergleich Schlaf/Nacht vs. Wach/Nacht sollte dagegen getestet werden, ob sich die überlegene Behaltensleistung durch Schlaf auch gegenüber einer Gruppe mit gleicher circadianer Phasenlage ergibt.

Das Lernmaterial war durch eine Paarassoziationsliste mit 16 Wortpaaren gegeben. Für diese Auswahl waren zwei Argumente maßgeblich. Einerseits zählen Paarassoziationslisten zu den Standard-Lernaufgaben der kognitiven (Schlaf-)Forschung, so dass hierdurch keine neue Variable eingeführt wurde, die eine Replizierbarkeit bereits bekannter Befunde erschwert oder verhindert hätte. Darüber hinaus eignen sich Wortpaar-Listen gut für die Anwendung wiederholter Abruf-Trials und ermöglichten so die Untersuchung der zweiten Fragestellung. Nachdem die Probanden die Liste in vier Study-Test-Durchgängen gelernt hatten, erfolgten am Ende des Retentionsintervalls anstelle des üblichen Einzeltests vier aufeinander folgende Test-Durchgänge, in denen jeweils nur die Stimulus-Worte präsentiert wurden und durch die zugehörigen Response-Worte ergänzt werden mussten.

Mit der Anzahl richtig ergänzter Wortpaare im letzten Testtrial der Lernsitzung als Ausgangsniveau zeigt Abb. 3 den prozentualen Verlust nach jeweils sieben Stunden im ersten Testtrial sowie kumuliert über alle vier konsekutiven Testtrials der Abrufsitzung. Kumuliertes Auszählen betrifft die Vereinigungsmenge der Treffer über alle vier Durchgänge und bedeutet damit, dass ein Wortpaar bei seiner ersten korrekten Komplettierung einmalig als Treffer gezählt wird, unabhängig davon, ob es in folgenden Trials erneut oder nicht mehr reproduziert wird. Während die Behaltensleistung in beiden Maßen in der Wach/Tag-Gruppe signifikant stärker nachgab als in der Schlafbedingung [einfacher Abruf: $t_d(37) = 2.19$; $p = .031$; kumulierter Abruf: $t_d(37) = 2.09$; $p = .039$],[2] zeigte sich zwischen den beiden Nachtbedingungen kein entsprechender Unterschied [einfacher Abruf und kumulierter Abruf jeweils $t_d(37) < 1$]. Gegenüber dem einfachen Abruf erbrachte die wiederholte Abrufprozedur in allen drei Gruppen einen vergleichbaren Rückgang des ursprünglichen Verlustes von knapp 10 % (Wach/Tag: 9.97 %; Wach/Nacht: 8.19 %; Schlaf/Nacht: 8.09 %), so dass sich das grundsätzliche Bild des Gruppenvergleiches hierdurch nicht änderte.

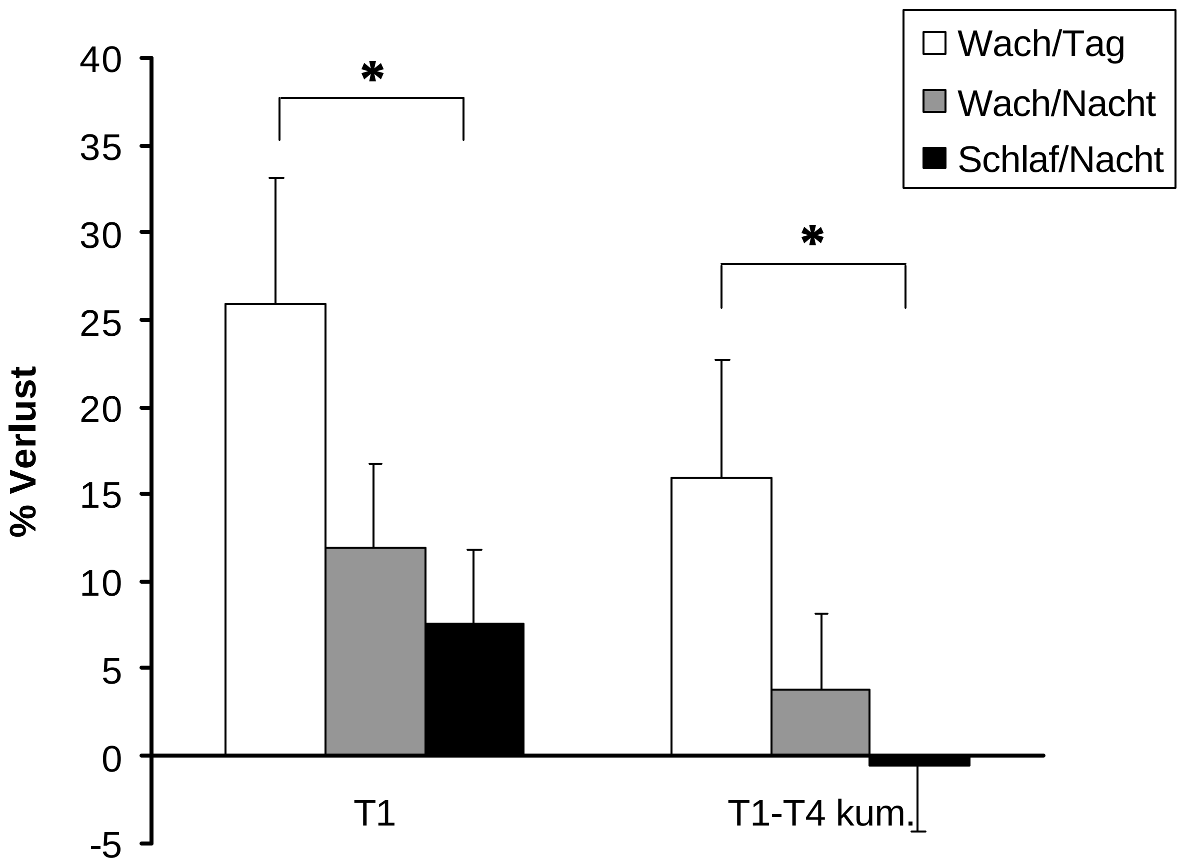


Abb. 3: Prozentualer Verlust an Wortpaar-Assoziationen (Mittelwert ± Standardfehler) bezogen auf das Ausgangsniveau am Ende der Lernsitzung im ersten Test-Trial (T1) und kumuliert über alle vier Test-Trials (T1-T4 kum.) nach sieben Stunden Wachheit während des Tages, Wachheit während der Nacht oder Schlaf während der Nacht. Aus: Lahl und Pietrowsky (eingereicht a). * p < .05.

[2] t_d bezeichnet die Dunnett *t*-Statistik für den Vergleich mehrerer Experimentalgruppen mit derselben Kontrollgruppe (hier der Schlafgruppe) (Dunnett, 1955).

Diese Ergebnisse legen nahe, dass im Zusammenhang mit dem „Schlaf-Effekt“ möglicherweise nicht der Schlaf selbst, sondern circadiane Faktoren der Nachtzeit eine entscheidende Rolle spielen. Offenbar leistet eine wache Periode während derselben circadianen Phase einen vergleichbaren Schutz des aufgenommenen Materials wie der Schlaf. Unterschiedliche Zugriffsmöglichkeiten während der Wiedergabe scheinen dabei keine Rolle zu spielen, wie sich an den vergleichbaren Zuwachsraten im Verlauf der Abrufwiederholungen zeigt. Es ist zu vermuten, dass neuroendokrine Faktoren, deren Konzentration unabhängig von Wachen und Schlafen nur mit der Tageszeit variiert, den Effekt vermitteln. In diesem Zusammenhang wird das Hormon Cortisol diskutiert, das beim Menschen offenbar hemmend auf das deklarative Gedächtnis wirkt (Kirschbaum, Wolf, May, Wippich & Hellhammer, 1996; Newcomer, Craft, Hershey, Askins & Bardgett, 1994; Wolkowitz et al., 1990) und dessen Freisetzung während des Schlafes der ersten Nachthälfte ein circadianes Minimum erreicht (Born et al., 1986; Weitzman et al., 1971). Allerdings wirkt Schlafdeprivation dieser Hemmung der Freisetzung entgegen (Von Treuer, Norman & Armstrong, 1996; Weitzman, Zimmerman, Czeisler & Ronda, 1983) – die Cortisolkonzentration variiert also nicht unabhängig vom Schlaf – so dass die weitgehend ungehinderte Gedächtnisleistung der Wach/Nacht-Gruppe nicht ohne weiteres auf verringerte Cortisol-Spiegel zurückgeführt werden kann.

7. Experiment II: Der Langzeiteffekt von Schlaf auf das Gedächtnis

Die Problemstellung dieser Arbeit betraf die zeitliche Stabilität des Schlafeffektes. Wenn unmittelbar im Anschluss an gleich lange Schlaf- und Wachepisoden eine bessere Behaltensleistung nach Schlaf gezeigt werden kann, bleibt diese Überlegenheit auch über längere Zeiträume erhalten? Die Beantwortung dieser Fragestellung könnte wichtige Hinweise im Hinblick auf funktionelle Aspekte des Schlafes liefern. Der Nachweis einer zeitlichen Robustheit des Effektes würde die Annahme stützen, dass die globale und weit reichende Optimierung der Gedächtnisperformanz eine zentrale Funktion des Schlafes darstellt. Eine nur flüchtige Erhöhung der Verfügbarkeit aufgenommener Inhalte, die nach kurzer Zeit wieder auf das Ausgangsniveau sinkt, wäre hingegen mit der Annahme eines bloßen Nebeneffektes besser verträglich.

Gemäß diesen Vorüberlegungen wurde neben dem Faktor Schlaf (Wach/Tag oder Schlaf/Nacht im Anschluss an das Lernen) auch die Dauer des Retentionsintervalls (7 h oder 72 h nach dem Lernen) in einem 2 × 2 Design variiert. Das Lernmaterial war eine Liste von 56 Substantiven, von denen immer vier Worte je einer von 14 semantischen Kategorien angehörten. In der Lernphase wurde diese Liste zweimal hintereinander in jeweils unterschiedlichen, aber für alle Probanden fixen Zufallsfolgen ohne Angabe der Kategorien präsentiert. Die verzögerte Abfrage beinhaltete zwei aufeinander folgende Abruftests mit einer Dauer von je 10 Minuten. Im ersten Durchgang (freier Abruf) waren die Versuchspersonen aufgefordert, möglichst viele der gelernten Worte in beliebiger Folge frei zu reproduzieren. Im zweiten Durchgang (gebundener Abruf) wurden die 14 Kategorien als Abrufreize simultan dargeboten und mussten mit den vier zugehörigen Instanzen gefüllt werden.

In der Sichtung der Ergebnisse stellte sich heraus, dass freier und gebundener Abruf zu $r = .95$ miteinander korrelierten und somit überwiegend dasselbe Konstrukt erfassten. Daher beschränkt sich die folgende Darstellung auf die Anzahl der Worte, die im freien Abruf des ersten Tests reproduziert werden konnte. Abb. 4 zeigt im Vergleich mit der Wachgruppe zunächst eine deskriptiv höhere Abrufleistung der Schlafgruppe nach Ablauf sowohl des kurzen als auch des langen Behaltenszeitraums. Während sich dieses Ergebnis für den 7 h Zeitraum inferenzstatistisch nicht erhärten ließ [$t(56) < 1$], deutete sich der Effekt für das 72 h Intervall an [$t(56) = 1.49$; $p = .071$] oder konnte sich sogar durchsetzen, wenn als Schätzer der Fehlervarianz nicht die Gesamtfehlerstreuung über alle vier Gruppen (*MSE*), sondern nur die mittlere Fehlerstreuung der beiden verglichenen Gruppen (Schlaf/72 h; Wach/72 h) verwendet wurde [$t(28) = 1.71$; $p = .049$].

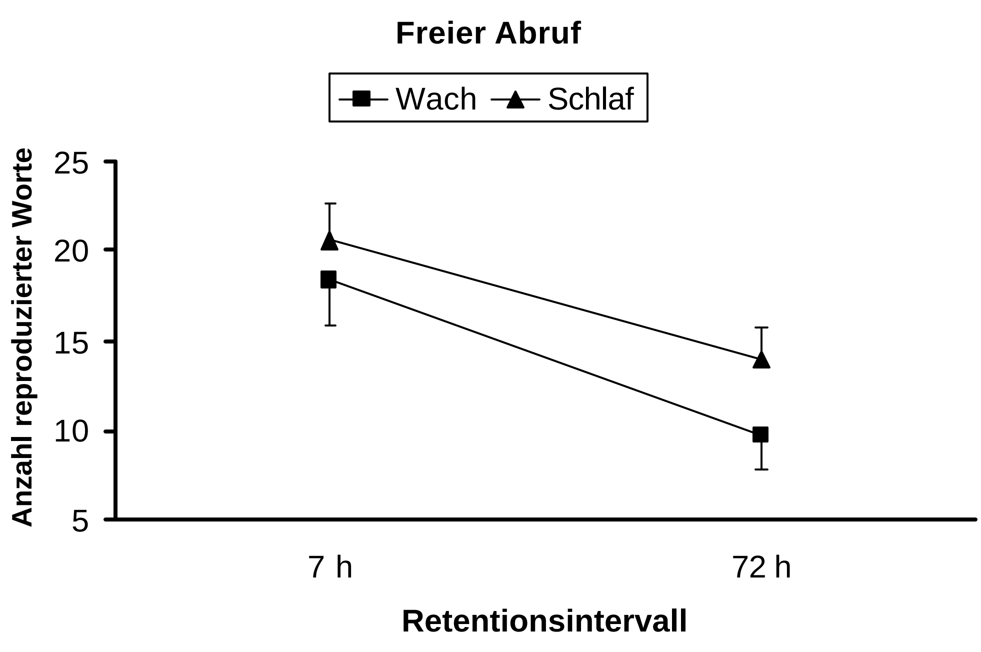


Abb. 4: Anzahl frei erinnerter Worte (Mittelwert ± Standardfehler) in Abhängigkeit von Wachheit oder Schlaf nach dem Lernen und der Dauer des Retentionsintervalls. Aus: Lahl und Pietrowsky (eingereicht b).

Nachdem sich in Experiment I eine bessere Paarassoziationsbildung der 7 h Schlaf/Nacht- gegenüber der 7 h Wach/Tag-Gruppe gezeigt hatte, ist die hier fehlgeschlagene Replikation für die freie Reproduktion von Worten unter der 7 h Bedingung als unerwartet zu werten und legt zunächst die Diskussion einer Materialabhängigkeit des Schlafeffektes nahe, wie sie etwa von Walker und Stickgold (2004) geführt wird. Dieser Annahme steht jedoch das Ergebnis für das 72 h Intervall gegenüber, in der dasselbe Material dann offenbar doch ausreichend sensitiv war, um auf Schlaf zu reagieren. Möglicherweise liegt eine komplizierte Wechselwirkung zwischen Material und Dauer des Behaltensintervalls vor.

Im Hinblick auf die ursprüngliche Fragestellung nach der zeitlichen Stabilität des Schlafeffektes lassen die Ergebnisse keine direkte Schlussfolgerung zu. Dazu wäre ein Wirksamkeitsnachweis für das kurze Zeitintervall erforderlich gewesen. In Analogie zu den Resultaten von Richardson und Gough (1963) – die Autoren fanden einen Schlafeffekt nach 144 h, nicht aber nach 24 h und 48 h – wirft das Experiment vielmehr die Frage auf „as to why the presence or absence of 8 hours sleep immediately after learning should in general aid retention […] over periods of up to 144 hours" (Richardson & Gough, 1963, S. 40). Eine theoretisch fundierte Erklärung für eine Wirksamkeitsverzögerung des Schlafeffektes von mehreren Tagen steht bislang aus. Im Sinne der Konsolidierungstheorie könnte man aber vermuten, dass Schlaf im Anschluss an das Lernen zwar eine günstige Voraussetzung für die Festigung des gelernten Materials ist, dass der gesamte Prozess der Konsolidierung aber erst deutlich später zum Abschluss kommt und damit auch erst verzögert wirksam wird. Unklar bleibt dann allerdings, warum der Zeitpunkt des Wirksamwerdens je nach Untersuchung deutlich variiert.

8. Experiment III: Begünstigt Schlaf das Vergessen irrelevanter Information?

Aus unserer Alltagserfahrung wissen wir, dass wir ständig eine Vielzahl an Informationen aufnehmen, die im weiteren Verlauf keine Bedeutung für uns oder die Bewältigung unserer täglichen Aufgaben haben. Das Vergessen irrelevanter oder irrelevant gewordener (obsoleter) Inhalte ist daher eine ebenso bedeutsame Funktion des Gedächtnissystems wie das Behalten relevanter Informationen (vgl. Zacks & Hasher, 1994). Die theoretische Grundlage von Experiment III war daher durch folgende Überlegung gegeben: Wenn sich Schlaf als ein Zustand etabliert hat, der maßgeblich an der Optimierung von Gedächtnisprozessen beteiligt ist, dann sollten auch adaptive Vergessensprozesse durch Schlaf begünstigt werden. In diesem Fall wäre also erstmalig eine schwächere Behaltensleistung nach Schlaf als nach

Wachsein zu erwarten. Wenn umgekehrt die Relevanz des Materials keinen Einfluss hätte, wäre die „klassische“ Wirkung von Schlaf, die Gedächtnisverstärkung, zu erwarten. Damit war das leitende Problem dieses Experiments eine zweiseitige Fragestellung, in der beide Richtungen eines möglichen Effektes von Interesse waren.

Die unabhängige Variable Behaltensintervall wurde analog zu Experiment I durch die drei Versuchsbedingungen Wach/Tag, Schlaf/Nacht und Wach/Nacht variiert. Die Operationalisierung irrelevanten Materials erfolgte durch zwei inzidentelle Lernaufgaben. In der ersten Aufgabe erhielten die Probanden eine Sequenz von 16 semantisch nicht assoziierten Wortpaaren. Unter dem Vorwand eines Kreativitätstests galt es, zu jedem Paar einen sinnvollen Satz zu bilden. Mit dieser Vorgabe wurde die zeitliche Begrenzung der Aufgabenrelevanz sichergestellt: Da der scheinbare Test mit dem erfolgreichen Lösen aller Aufgaben beendet war, wurde das Material damit im Rahmen des Experiments obsolet, und so gab es aus Sicht der Probanden keine Veranlassung mehr, sich weiterhin mit dem Stoff zu beschäftigen oder gar eine tiefere Einprägung zu versuchen. Am Ende des Retentionsintervalls wurde jedoch überraschend die Behaltensleistung für die solcherart inzidentell gelernten Wortpaare durch drei aufeinander folgende gebundene Abruf-Durchgänge gemessen. Nur noch die Stimulus-Wörter eines jeden Paares wurden dabei sequenziell präsentiert und mussten durch die zugehörigen Response-Wörter komplettiert werden.

Im Rahmen der zweiten Aufgabe wurde eine Sequenz von 24 affektiv neutralen Alltagsgegenständen präsentiert. Hierbei galt nun die (Fehl-)Instruktion, „sich so viele Objekte wie möglich einzuprägen.“ Die Bilder wurden indes nicht in der Mitte, sondern jeweils an einer der vier möglichen Eckpositionen des Bildschirms präsentiert. Am Ende des Behaltensintervalls wurden die Bilder erneut – nunmehr in der Mitte des Bildschirms – gezeigt und zu jedem Bild musste angegeben werden, in welchem Quadranten es während der Lernphase präsentiert worden war. Wiederum wurde also durch die Art der Präsentation und den Einsatz einer fehlleitenden Instruktion die beiläufige Form der Akquisition sichergestellt.

Am Ende des Versuches wurden alle Teilnehmer hinsichtlich einer möglichen Antizipation der beiden unangekündigten Gedächtnistests befragt. Zwei Probanden der Schlaf/Nacht-Gruppe und drei Probanden der Wach/Tag-Gruppe gaben daraufhin an, den Abruf der Paarassoziationen bereits während des Lernens erwartet zu haben und wurden daher von der Analyse dieser Aufgabe ausgeschlossen. Die Abfrage der Bildpositionen war hingegen von keinem der Teilnehmer erwartet worden. Der insgesamt sehr niedrige Anteil an Versuchspersonen, die die Natur des Experiments im Vorfeld durchschaut hatten, unterstreicht die gelungene Umsetzung des inzidentellen Lernens.

Abb. 5 zeigt die kumulierte Anzahl korrekt komplettierter Wortpaare über die drei aufeinander folgenden Abruf-Durchgänge. Die höhere Abrufleistung der Schlafgruppe gegenüber den beiden Wachgruppen verfehlte jeweils knapp die inferenzstatistische Absicherung [Schlaf / Nacht vs. Wach / Nacht: $t_d(47) = 2.22$; $p = .057$; Schlaf/ Nacht vs. Wach / Tag: $t_d(47) = 1.94$; $p = .10$; zweiseitige *p*-Werte].

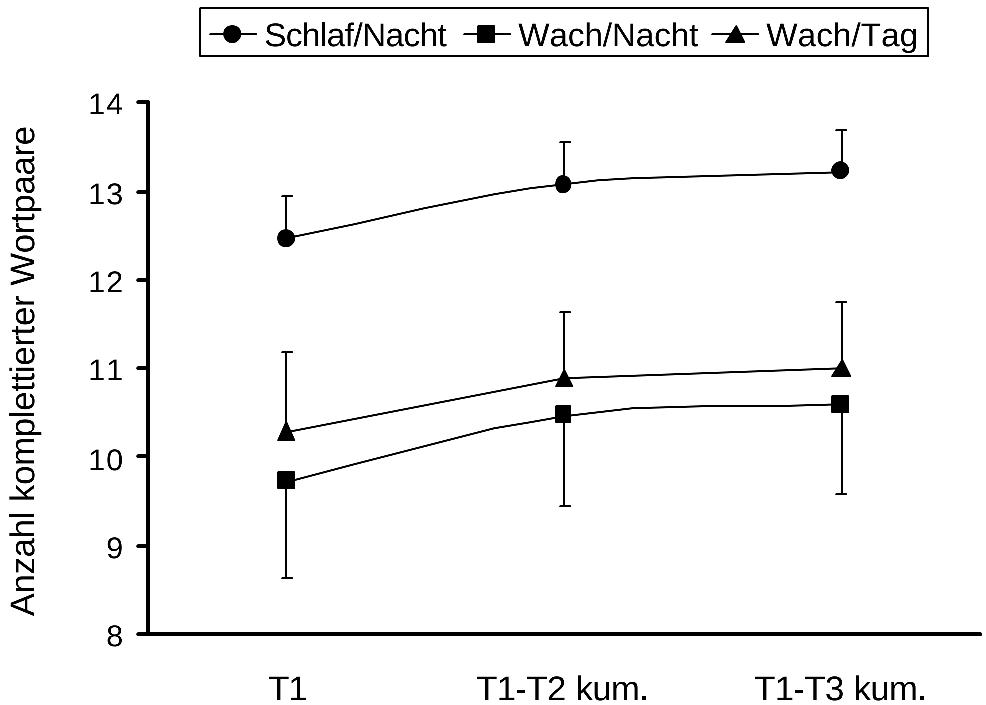


Abb. 5: Kumulierte Anzahl inzidentell gelernter Wortpaare (Mittelwert ± Standardfehler), die nach sieben Stunden Schlaf, Wachheit während der Nacht oder Wachheit während des Tages korrekt komplettiert wurden. Aus: Lahl und Pietrowsky (eingereicht c).

In Abb. 6 ist die Anzahl richtig zugeordneter Bildpositionen wiedergegeben. Es fällt zunächst auf, dass sich die Erinnerungsleistung in allen drei Gruppen deutlich über dem binomialen Zufallsniveau (10 Positionen) bewegte.[3] Ähnlich wie im Falle des inzidentellen Paarassoziationslernens ergab sich unter der Schlafbedingung eine höhere Anzahl richtig zugeordneter Bildpositionen als in beiden Wachgruppen. Dieser Unterschied erwies sich als statistisch signifikant [$t_d(47) = 2.78$; $p = .015$] für den

[3] Bei einer Sequenz von $n = 24$ (als voneinander unabhängig angenommenen) Bildern und einer Ratewahrscheinlichkeit von $p = ¼$ je Bild folgt die Summe S der zufällig richtig zugeordneten Positionen einer $B(24; ¼)$-Verteilung, deren 5 % Quantil 10 Bilder sind: $P(S \geq 10) < 0.05$.

Vergleich Schlaf/Nacht vs. Wach/Nacht, nicht aber für den Vergleich Schlaf/Nacht vs. Wach/Tag [$t_d(47) = 1.36$; $p = .30$; zweiseitige p-Werte].

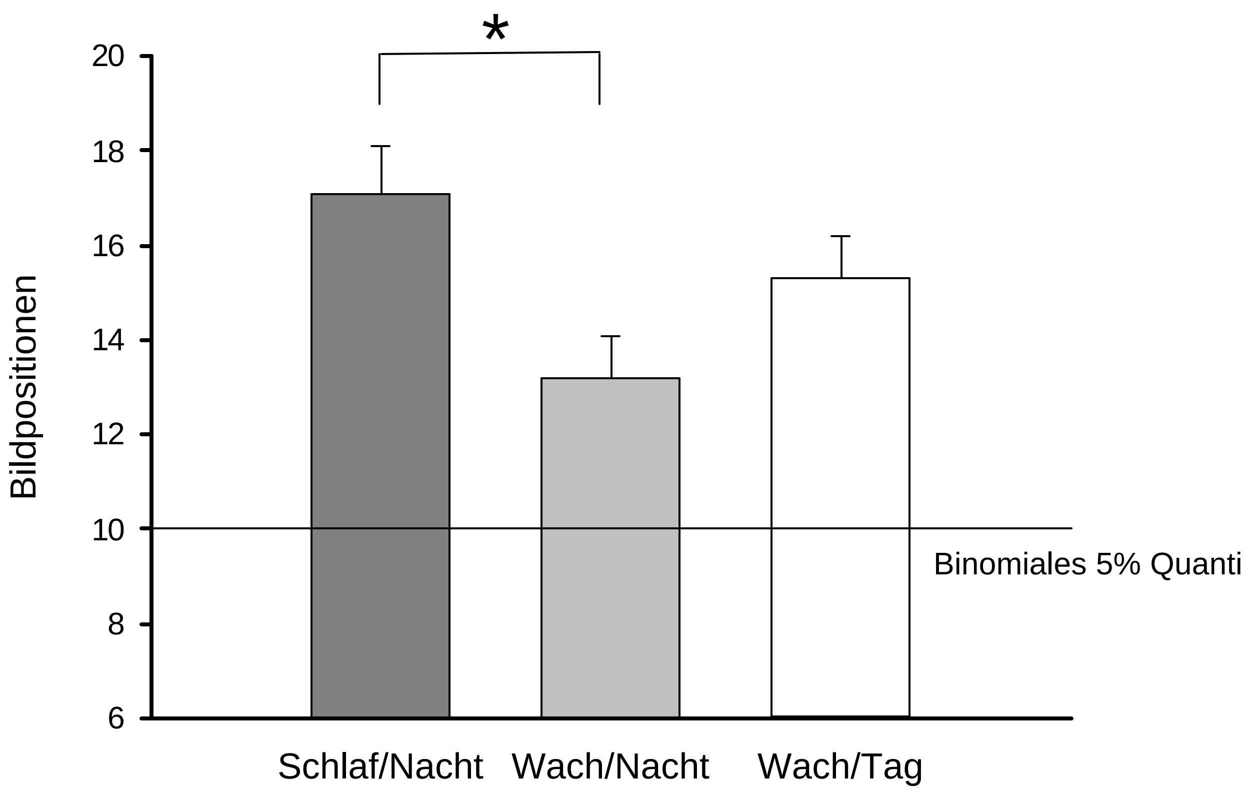


Abb. 6: Anzahl inzidentell gelernter Bildpositionen (Mittelwert ± Standardfehler), die nach sieben Stunden Schlaf während der Nacht, Wachheit während der Nacht oder Wachheit während des Tages richtig zugeordnet werden konnten. Eingezeichnet ist auch das binomiale 5 % Zufallsniveau, das von allen drei Gruppen deutlich überschritten wurde. Aus: Lahl und Pietrowsky (eingereicht c). * $p < .05$.

Mit diesen Befunden lässt sich die Hypothese einer verbesserten Entfernung irrelevanter Inhalte durch Schlaf nicht bestätigen. Im Gegensatz zur Annahme einer solchen Filterfunktion deuten die Ergebnisse sogar eine verbesserte Abrufleistung nach Schlaf an, auch wenn Beeinträchtigungen durch Müdigkeit infolge von Schlafdeprivation hier eine große Rolle zu spielen scheinen. Zusammenfassend ergeben sich aus diesem Experiment keinerlei Hinweise auf eine Spezifität des Schlafeffektes im Hinblick auf die Bedeutsamkeit des aufgenommenen Materials.

9. Experiment IV: Der Effekt von Schlaf auf falsche Erinnerungen

In diesem Versuch wurde der übliche Ansatz der Behaltenmessung fallengelassen. Statt die Erinnerung an Material zu testen, das zuvor präsentiert wurde, ging es hier um fehlerhafte Erinnerungen an *nicht* präsentierte Inhalte, mithin um falsche Erinnerungen. Obwohl die experimentelle Untersuchung falscher Erinnerungen einerseits und die Erforschung des Gedächtnisses im Schlaf andererseits in den vergangenen Jahren stark zugenommen hat (vgl. Bruce & Winograd, 1998; Stickgold & Walker, 2005), fanden sich nach Sichtung der Literatur keine Arbeiten, in denen beide Bereiche zusammengeführt worden wären. In diesem Experiment wurde damit erstmalig die Frage untersucht, ob es einen Schlaf-Effekt für falsche Erinnerungen gibt. In Ermangelung bestehender Ergebnisse, auf die sich eine klare theoretische Vorhersage hätte beziehen können, lag dem Experiment eine unspezifische (zweiseitige) Fragestellung zugrunde.

In einer einflussreichen Arbeit haben Roediger und McDermott (1995) ein einfaches Instrument zur Induktion falscher Erinnerungen im Labor vorgestellt. Probanden hören Wortlisten (z.B. „weiß, dunkel, verkohlt, Beerdigung, …“), deren Elemente alle semantisch mit einem kritischen Konzeptwort („schwarz“) assoziiert sind, das seinerseits *nicht* genannt wird. Interessanterweise neigen Versuchspersonen in folgenden Abruftests nicht nur dazu, diese nicht präsentierten „Köder“ irrtümlicherweise wieder zu erkennen oder frei zu reproduzieren, sie tun dies auch mit großer subjektiver Überzeugung (Roediger & McDermott, 1995; Toglia, Neuschatz & Goodwin, 1999).

Zwei Versuchsgruppen lernten 18 derartige Wortlisten in einer einmaligen akustischen Präsentation und waren nach sieben Stunden Schlaf in der Nacht oder Wachheit während des Tages aufgefordert, je Liste sechs Worte als „alt“ (zuvor präsentiert) oder „neu“ (zuvor nicht präsentiert) zu klassifizieren. Neben je zwei nicht präsentierten nicht-assoziierten Kontrollworten waren von diesen sechs Worten je drei während des Lernens dargeboten worden (Listen-Worte) und erfassten so die richtigen Erinnerungen. Je ein weiteres Wort war durch das nicht genannte Ziel- oder Köder-Wort gegeben und diente somit der Bestimmung der falschen Erinnerungen.

Abb. 7 zeigt den Prozentsatz der Listen-Worte (richtige Erinnerungen) und der Köder-Worte (falsche Erinnerungen), die unter den beiden Versuchsbedingungen als „alt“ klassifiziert (wieder erkannt) wurden. Während Schlaf keinen Effekt auf das richtige Wiedererkennen hatte [$t(24) < 1$], zeigte die Schlafbedingung gegenüber der Wachgruppe eine drastisch erhöhte Anzahl falscher Erinnerungen [$t(24) = 4.70$; $p < .0001$; zweiseitiger p-Wert].

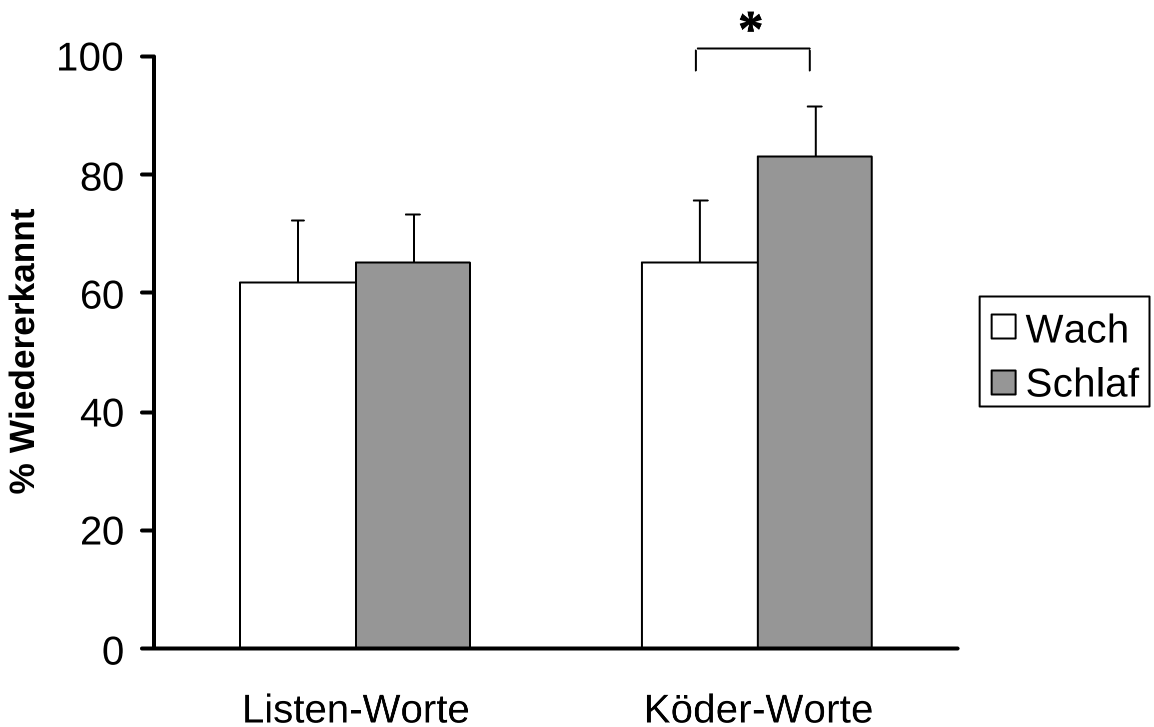


Abb. 7: Relative Anzahl (Mittelwert ± Standardabweichung) wieder erkannter Listen-Worte (richtige Erinnerungen) und Köder-Worte (falsche Erinnerungen) nach sieben Stunden Wachheit oder Schlaf. Nach Daten aus: Lahl et al. (eingereicht). * $p < .05$.

Die erhaltenen Daten überraschen in zweierlei Hinsicht. Zunächst wäre auf der Grundlage der bestehenden Literatur ein positiver Effekt von Schlaf auf die richtigen Rekognitionen zu erwarten gewesen, wie er etwa von Nesca und Koulack (1994) mit zwei vergleichbaren Versuchsbedingungen demonstriert werden konnte. Darüber hinaus trat die schlafbedingte Verstärkung der falschen Erinnerungen mit einer unerwartet hohen Effektstärke von nahezu zwei Standardabweichungen (Cohen's d = 1.84) in Erscheinung. Möglicherweise lag die Ursache für beide Phänomene in der sehr speziellen Natur der hier verwendeten Wortlisten. Denkbar wäre etwa, dass ein etwaiger Effekt von Schlaf auf die richtigen Erinnerungen durch die falschen Erinnerungen überschattet wurde.

Die Ergebnisse stehen indes in guter Übereinstimmung mit Befunden anderer Arbeiten, in denen gezeigt werden konnte, dass richtige und falsche Erinnerungen differenziell auf Manipulationen sowohl des Retentionsintervalls (Payne, Elie, Blackwell & Neuschatz, 1996; Seamon et al., 2002) als auch der Prozesse des Enkodierens und Abrufens (Huron, Servais & Danoin, 2001) reagieren. Zusammen mit den hier vorgestellten Resultaten verdichten sich damit die Hinweise auf grundsätzlich unterschiedliche interne Repräsentationen (Engramme) korrekter und fehlerhafter Speicherinhalte im episodischen Gedächtnis.

Als theoretischer Erklärungsrahmen bietet sich hier unmittelbar die Fuzzy Trace Theory (Brainerd & Reyna, 2002; Brainerd, Reyna & Poole, 2000) an, besteht doch ihr inhaltlicher Kern in der Annahme eines dualen Speicherprozesses. Demnach werden Elemente im Gedächtnis in Form zweier paralleler, dissoziierter Spuren abgelegt. Die so genannte wortgetreue Spur (engl. „verbatim trace”) erfasst die konkreten Eigenschaften und Details eines Elementes (m. a. W. die Instanz, z. B. das Holz, die Anzahl der Beine, die Größe eines Tisches). Dagegen repräsentiert die sinngemäße Spur (engl. „gist trace“) die übergeordnete Bedeutung, das gemeinsame Konzept (m. a. W. die Kategorie, Möbel). Im Zusammenhang mit falschen Erinnerungen bilden die assoziierten Listen-Worte die Grundlage der wortgetreuen Spur, wohingegen die nicht genannten Konzeptworte der sinngemäßen Spur entsprechen.

10. Methoden: Die automatische Genenerierung paralleler Wortlisten

Die kognitive Schlafforschung gehört im Hinblick auf die Versuchsdurchführung zu den aufwändigen Teilbereichen der Psychologie. Dort, wo andere kognitive Disziplinen im Extremfall einen kompletten Datensatz innerhalb weniger Minuten im Hörsaal erheben können (z. B. Nilsson, Law & Tulving, 1988; Roediger & McDermott, 1995), fällt im Schlaflabor für jeden Probanden ein erheblicher apparativer und zeitlicher Aufwand von 10 Stunden und mehr pro Versuchsnacht an. Stärker als in anderen Bereichen beeinflusst damit die Versuchsökonomie die experimentelle Planung und mahnt zur Verwendung von Designs, die bei möglichst geringer Fallzahl maximale Teststärke (Power) aufweisen.

An dieser Stelle bietet sich die Anwendung von Messwiederholungsplänen an. Die Kontroverse um mögliche asymmetrische Transfereffekte zwischen zwei Behandlungen und weitere Nachteile dieser Versuchspläne (Krauth, 2000; McGuigan, 1978; Namboodiri, 1972) soll hier nicht aufgegriffen werden. Jedoch kann ein abhängiges Design (bei gleicher totaler Anzahl der Messungen) eine höhere statistische Power aufweisen als ein entsprechendes unabhängiges Design, wenn die wiederholten Messungen hoch miteinander korreliert sind, wenn also die intraindividuelle Fehlervarianz deutlich geringer ausfällt als die interindividuelle Fehlervarianz (Howell, 2002).

Wenn die Testung der verbalen Gedächtnisleistung wiederholt an einer Person erfolgen soll, muss in jeder Messung semantisch unterschiedliches Wortmaterial verwendet werden, um direkte Lerneffekte zu vermeiden. Im Widerspruch dazu fordert das Prinzip der isolierenden Variation der unabhängigen Variable(n) eine möglichst perfekt übereinstimmende Lern- und Erinnerbarkeit der unterschiedlichen Listen. Die kognitive Psychologie hat eine Reihe psycholinguistischer Variablen bestimmt, die Einfluss auf die Enkodierung, Retention und den Abruf von Worten haben können. Neben der Wortlänge und der Worthäufigkeit in der geschriebenen und gesprochenen Sprache dürften die Skalen Konkretheit, Bildhaftigkeit und Bedeutungshaltigkeit von Paivio (Paivio, Yuille & Madigan, 1968) sowie die Osgood'schen Dimensionen der Valenz, Potenz und Erregung (Osgood & Suci, 1955) zu den geläufigsten Attributen zählen. Sollen allgemein *n* psycholinguistische Attribute Berücksichtigung finden, so kann formal jedes Wort einer gegebenen Wortmenge als Vektor in einem *n*-dimensionalen Eigenschaftsraum repräsentiert werden. Die Bestimmung der idealen Wortlisten für eine zweimalige Messung erfordert dann das Auffinden derjenigen Vektorpaare, die den geringst möglichen Abstand zueinander aufweisen. Bei $k > 2$ Messungen muss entsprechend die Menge der *k*-Tupel mit dem geringsten Abstand gefunden werden.

Die vom Autor entwickelte Software *EquiWord* leistet genau diese Berechnungen für die gebräuchlichsten Distanzkoeffizienten. Experimentatoren können dabei problemlos auf bestehende Datenbanken mit Wortnormen zurückgreifen. Neben generischen Datenbankformaten (Excel, comma separated value files) ist das Programm auch in der Lage, das Datenformat der englischsprachigen MRC-Datenbank (Wilson, 1988) zu lesen, die mit einem Pool von insgesamt 150837 Worten und zugehörigen Ratings auf bis zu 26 Dimensionen das umfangreichste psycholinguistische Datenmaterial in elektronisch lesbarer Form stellen dürfte. Darüber hinaus hat der Autor die von Hager und Hasselhorn (1994) zur Verfügung gestellten deutschen Norm-Dateien mit den Häufigkeitsdaten der deutschsprachigen Celex-Datenbank (Baayen, Piepenbrock & Gulikers, 1995) zu einer gemeinsamen Datenbank zusammengeführt, die vom Programm gelesen werden kann, so dass auch für deutschsprachige Untersuchungen ein Pool von 1814 Worten (750 Substantive, 448 Verben, 616 Adjektive) mit je acht Attributen für die Erstellung paralleler Wortlisten zur Verfügung steht.[4] Kürzlich haben z. B. auch Izura, Hernández-Muñoz und Ellis (2005) Wortnormen für 500 spanische Worte auf je sieben psycholinguistischen Attributen in maschinenlesbarer Form veröffentlicht, die somit ebenfalls für die Verarbeitung mit *EquiWord* zur Verfügung stehen.

Tab. 1 zeigt beispielhaft eine Wortpaar-Liste, die mit *EquiWord* aus der Teilmenge der 616 verfügbaren deutschen Adjektive erstellt wurde. Aufgeführt sind die 15 Adjektivpaare mit den geringsten Mahalanobis-Abständen aus der Gesamtmenge der $(616^2 - 616) / 2 = 189420$ möglichen Paare. Wie

[4] Die Software kam den eigenen Untersuchungen indes nicht mehr zugute, da sie einerseits erst nach Durchführung der Experimente I und II fertig gestellt wurde und andererseits inzidentelle und falsche Erinnerungen (Experimente III und IV) naturgemäß nur einmal je Person gemessen werden können.

man sieht, korrespondieren die metrischen mitunter mit den semantischen Nachbarschaften („platt-schmal“, „jähzornig-rüpelhaft“). Das muss aber nicht so sein („biologisch-reparabel“, „funktionell-gönnerisch“).

Tab. 1: Beispiel für eine mit *EquiWord* generierte Liste von psycholinguistisch eng benachbarten Wortpaaren.

Adjektiv 1	Adjektiv 2	Mahalanobis-Abstand
geschickt	großzügig	0.464
loyal	nobel	0.625
biologisch	reparabel	0.648
platt	schmal	0.651
gesprächig	rhythmisch	0.655
munter	spaßig	0.658
seltsam	typisch	0.663
jähzornig	rüpelhaft	0.664
vernünftig	vorbildlich	0.680
bockig	gierig	0.684
ernsthaft	gründlich	0.693
äußerlich	nachlässig	0.700
komisch	locker	0.703
ähnlich	willig	0.704
funktionell	gönnerisch	0.727

Sollen mehr als zwei, also allgemein k Bedingungen wiederholt an denselben Probanden getestet werden, so erfordert dies die Erstellung von k äquivalenten Listen. Sind aber aus einer Menge von Worten jeweils alle möglichen k-Tupel zu bilden, so steigt die Anzahl der möglichen Kombinationen und damit der Rechenaufwand rasch an. So fallen im oben genannten Beispiel der 616 Adjektive bereits knapp sechs Milliarden mögliche Kombinationen an, wenn anstelle der benachbarten Paare benachbarte Quadrupel gebildet werden sollen, und ein amerikanischer Forscher, der dasselbe mit der viel umfangreicheren MRC-Datenbank bewerkstelligen wollte, müsste in Anbetracht von etwa 2×10^{19} Kombinationen etliche Millionen Jahre auf das Ergebnis warten. *EquiWord* bietet für diese Szenarien ein alternatives Vorgehen:[5] Zu jedem Wort einer *vorgegebenen* Teilmenge von Worten können die $k-1$ engsten Nachbarn berechnet werden. Dieses Verfahren fordert zwar vom Benutzer die willkürliche

[5] Dieses Programm-Feature wurde erst nach Annahme der Publikation implementiert.

Vorgabe der Startwörter, der Rechenaufwand steigt dafür aber nur noch linear mit der Größe der Tupel. Tab. 2 illustriert das Feature für die ersten drei Adjektive der ersten Spalte aus Tab. 1 und die zugehörigen fünf Nachbarworte mit dem geringsten Mahalanobis-Abstand.

Tab. 2: Beispiele für drei Sextupel psycholinguistisch benachbarter Worte, die mit *EquiWord* berechnet wurden.

Startwort	Nachbarn	Mahalanobis-Abstand
geschickt		
	großzügig	0.464
	herzlich	0.913
	rhythmisch	1.166
	brillant	1.236
	zielbewusst	1.257
loyal		
	nobel	0.625
	ehrbar	0.696
	global	0.814
	artig	1.269
	seriös	1.293
biologisch		
	reparabel	0.648
	erfreulich	1.034
	umgänglich	1.116
	äußerlich	1.213
	selbstlos	1.231

11. Diskussion

In der Übersicht haben die hier vorgestellten Arbeiten gezeigt, dass der Effekt von Schlaf auf das Gedächtnis durch circadiane Faktoren mitbestimmt wird (Exp. I), verzögert einsetzen kann (Exp. II), kein Zugriffs- bzw. Abrufartefakt darstellt (Exp. I-III) und keine Filterfunktion bezüglich irrelevanter (Exp. III) oder falscher Erinnerungen (Exp. IV) vermittelt. Im Hinblick auf die eingangs aufgeworfene Frage nach der Generalisierbarkeit des Effektes von Schlaf auf das Gedächtnis geben diese Befunde keine ausreichende Evidenz für eine stabile und umfassende Optimierung der Speicherung und Filterung neu aufgenommener Informationen durch Schlaf. So war die erwartete positive Wirkung von Schlaf auf richtige Erinnerungen nicht nachweisbar, wenn die circadiane Phase konstant gehalten wurde (Exp. I), wenn kategoriale Wortlisten über ein kurzes Behaltensintervall untersucht wurden (Exp. II) oder wenn semantisch assoziierte („Köder-")Wortlisten als Testmaterial dienten (Exp. IV). Umgekehrt ergab sich anstelle einer Abschwächung ein tendenziell erhöhtes Behalten irrelevanter Inhalte (Exp. III) sowie eine deutliche Verstärkung falscher Erinnerungen (Exp. IV). Alle gemessenen Effekte sind indes unabhängig von Faktoren des Zugriffs. In keiner der wiederholten gebundenen Abrufprozeduren der Experimente I-III ergab sich ein Hinweis auf differenzielle Zugewinne der Schlaf- oder Wachgruppen.

Während also Faktoren der Tageszeit und Behaltensdauer den Zusammenhang zwischen Schlaf und Gedächtnis offenbar beträchtlich komplizieren, muss auf der anderen Seite festgestellt werden, dass es mit einer einzigen dem Verfasser bekannten Ausnahme (Portnoff, Baekeland, Goodenough, Karacan & Shapiro, 1966) keine Untersuchung gibt, die im Vergleich zwischen Wach und Schlaf je eine *schlechtere* Erinnerungsleistung nach Schlaf demonstriert hätte. Typischerweise setzt sich die Gedächtnis fördernde Wirkung von Schlaf entweder mit statistischer Signifikanz durch oder sie bleibt aus (z. B. Benson & Feinberg, 1975; Nesca & Koulack, 1994; Wagner, Gais & Born, 2001), sie kehrt sich aber nicht in eine Beeinträchtigung um. Die hier vorgestellten Untersuchungen bilden da keine Ausnahme. Sie demonstrieren mit den Experimenten III und IV sogar recht deutlich die mangelhafte Selektivität des Schlafeffektes in Bezug auf Materialen, die im Sinne der gestellten Aufgaben besser nicht behalten werden sollten. Die in Tab. 3 gezeigte Aufstellung der Cohen'schen Effektstärken illustriert die genannte Einseitigkeit der Effektrichtungen, alle Effektstärken haben positives Vorzeichen, aber auch die Varianz ihres Ausmaßes.

Tab. 3: Effektstärken (Cohens *d*) der paarweisen Gruppenvergleiche aus den vorgestellten Experimenten.

	Schlaf vs. Wach / Tag	Schlaf vs. Wach / Nacht
Exp. I: % Verlust Paarassoziationen		
einfacher Abruf	0.84	0.20
kumulierter Abruf	0.84	0.22
Exp. II: Anzahl Worte nach		
7 h	0.28	—
72 h	0.54	—
Exp. III: Anzahl inzidentell gelernter		
Paarassoziationen	0.71	0.84
Bildpositionen	0.46	1.01
Exp. IV: Anzahl Worte		
richtig	0.39	—
falsch	1.84	—

Beachte. Positive Werte zeigen eine bessere Gedächtnisleistung nach Schlaf an.

Wie lassen sich die Ergebnisse im Rahmen einer übergeordneten Theorie deuten? In der Summe legen die Befunde eine Erklärung im Sinne einer retrograden Erleichterung nahe, wie sie auch für Benzodiazepine (Coenen & Van Luijtelaar, 1997; Hinrichs, Ghoneim & Mewaldt, 1984) und Alkohol (Mueller, Lisman & Spear, 1983; Parker et al., 1980) diskutiert werden. Beide Agenzien wirken anterograd amnestisch, hemmen also die Enkodierung von Material, das nach der Einnahme dargeboten wird. Über diesen Mechanismus reduzieren sie aber gleichzeitig das Ausmaß retroaktiver Hemmung und wirken somit protektiv auf Inhalte, die *vor* der Einnahme aufgenommen wurden (Wixted, 2004). In ähnlicher Weise scheint Schlaf zuvor aufgenommenes Material durch die anterograde Blockierung neuer Inhalte zu schützen (Abb. 8). Dieser Erklärungsansatz entspricht der Interferenztheorie des Vergessens, wie sie ursprünglich von McGeoch (1932) und in jüngerer Zeit von Coenen und Van Luijtelaar (1997) (re)formuliert wurde. Da sie eher eine passive oder beiläufige Abschirmung des Gedächtnisses vor neu eintreffender Information beinhaltet, kann diese Theorie die gleichsinnigen, aber stark variierenden Effekte sowie die geringe Spezifität im Hinblick auf das Material besser erklären als die Annahme aktiver und spezifischer Konsolidierungsprozesse im Schlaf, von denen man eine größere Zuverlässigkeit und Selektivität erwartet hätte.

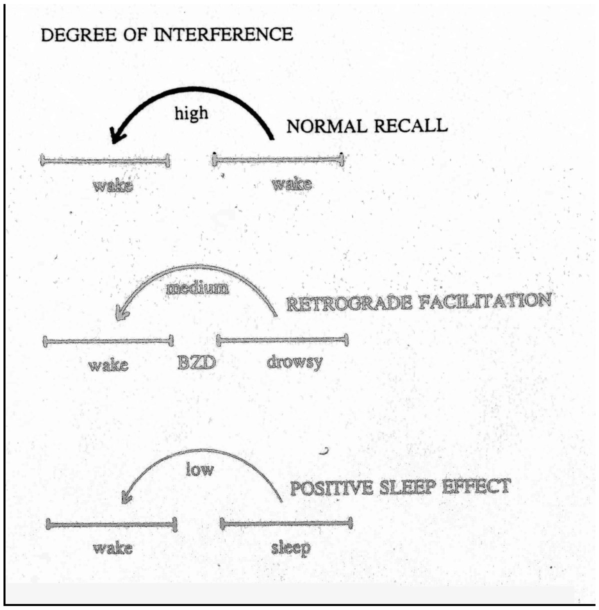


Abb. 8: Übergeordnetes Modell zur Erklärung der retrograden Erleichterung durch Benzodiazepine (BZD) und Schlaf. Der Schutz vor retrograder Hemmung wird als gemeinsamer Wirkmechanismus angesehen. Aus: Coenen und Van Luijtelaar (1997).

12. Ausblick

Die hier vorgestellten Experimente geben erste Hinweise darauf, dass der Effekt von Schlaf auf das Gedächtnis eher beiläufiger oder sekundärer, statt primär funktioneller Natur ist. Genau wie andere, alternative Theorien erfordert dieser Entwurf natürlich weitere empirische Bestätigungen, um künftig größeren Einfluss auf die Diskussion nehmen zu dürfen. So ist zu wünschen, dass nachfolgende Untersuchungen die Frage nach einer möglichen Filterung von Informationen aufgreifen und vertiefend behandeln. Hilfreich wäre etwa ein Experiment, in dem das Behalten relevanter *und* irrelevanter Inhalte gemeinsam getestet wird. Der Quotient beider Behaltensmengen wäre ein direkter Index etwaiger Selektionsprozesse (oder des Mangels derselbigen). Auch die überraschende Verstärkung falscher Erinnerungen im Schlaf erfordert eine Replikation, daneben aber auch Folgeexperimente, durch die dieses kontraintuitive Ergebnis einer theoretischen Erklärung zugänglich gemacht wird. Schließlich steht eine umfassende Klärung des Langzeit-Effektes weiterhin aus. Möglicherweise könnte hier eine größer angelegte Untersuchung über mehrere gestaffelte Retentionsintervalle Aufschluss geben. Das Experiment sollte dabei über genügend statistische Power verfügen, um gegebenenfalls auch kleine und mittlere Effekte aufzudecken. *EquiWord* würde hier für die Umsetzung eines sparsamen Messwiederholungsdesigns bereit stehen.

In Anbetracht des Ringens der Schlafforschung um die seit langem ausstehende Entschlüsselung der Funktion des Schlafes könnte man abschließend versucht sein, eine Analogie zur Funktion des Essens herzustellen. Der Umstand, dass Menschen und viele andere Säugetiere ihre Nahrung in Gemeinschaft verzehren, deutet auf eine wichtige Funktion des Essens für den Erhalt der sozialen Integrität hin. Zweifellos ist dieser Effekt aber von nachrangiger Bedeutung und wird an vielen Stellen durch die primäre Funktion des Stoffwechsels übersteuert. Ein außerirdischer Forscher, der seine Untersuchungen von der Hypothese leiten ließe, dass wir primär essen, um den sozialen Zusammenhalt zu sichern, würde sicherlich eine Vielzahl empirischer Betätigungen hierfür finden. Er würde aber auch auf eine Fülle heterogener oder schwacher Befunde stoßen, wann immer sich die eigentliche – ihm unbekannte – Funktion der Verstoffwechselung gegen die nach geschalteten Funktionen durchsetzte. Etwa, wenn ein Raubtier, das sonst seine Mahlzeiten teilt, nach einer langen Hungerperiode ohne Rücksicht auf seine Artgenossen alleine speist. Derzeit sieht es so aus, als befände sich die Schlafforschung in einer ähnlich unglücklichen Position wie der Außerirdische.

13. Zusammenfassende Thesen

- Die Funktion des Schlafes ist bislang ungeklärt. Viele Befunde deuten aber darauf hin, dass Schlaf eine günstige Wirkung auf die Gedächtnisleistung hat. Typischerweise zeigen Probanden, die zwischen Erwerb und Abruf eines bestimmten Lernmaterials geschlafen haben, eine bessere Behaltensleistung als Probanden, die während dieser Zeit wach blieben (sog. Schlafeffekt). Damit könnte die Optimierung von Gedächtnisfunktionen eine zentrale Aufgabe sein, die im Schlaf bewältigt wird.
- Unter dieser Annahme einer funktionellen Bedeutung des Schlafes für das Gedächtnis ist eine stabile und ausdifferenzierte Begünstigung verschiedener Gedächtnisleistungen durch Schlaf zu erwarten. Mit den hier vorgestellten Experimenten sollte diese Vorhersage auf unterschiedlichen Ebenen empirisch getestet werden.
- Im Einzelnen wurde geprüft, ob der beschriebene Effekt (i) tatsächlich durch Schlaf oder aber durch circadiane Faktoren der Nachtzeit vermittelt wird, (ii) von der Methode des Abrufs aus dem Langzeitgedächtnis abhängt, (iii) über längere Zeiträume erhalten bleibt und wie sich Schlaf auf (iv)

beiläufig gelernte Inhalte und (v) Trugerinnerungen auswirkt.

- Unter der Voraussetzung einer funktionellen Verknüpfung zwischen Schlaf und Gedächtnis war zu erwarten, dass (i) Schlaf (und nicht Nachtzeit) die entscheidende Variable für eine Begünstigung der Behaltensleistung ist, der Schlaf-Effekt unabhängig vom (ii) Abrufverfahren und der (iii) Dauer des Behaltensintervalls auftritt und dass (iv) irrelevante und (v) falsche Erinnerungen im Schlaf besser zurückgewiesen werden können als im Wachzustand.
- Die Ergebnisse zeigten, dass (i) die circadianen Faktoren der Nachtzeit auch unabhängig von Schlaf zu einer Verbesserung der Behaltensleistung führen können, dass (ii) unterschiedliche Methoden des Abrufs keine Rolle spielen, dass dieser Effekt (iii) mit zeitlicher Verzögerung einsetzen kann und (iv) teilweise auch auf irrelevante Erinnerungen und (v) deutlich auf falsche Erinnerungen generalisiert, aber (i, iii, v) für richtige Erinnerungen nicht unabhängig von anderen Variablen nachstellbar ist.
- Die empfindliche Abhängigkeit von anderen Faktoren sowie die mangelnde Selektivität in Bezug auf unerwünschte Inhalte sprechen gegen die Annahme spezifischer Prozesse, die funktionell an den Schlaf gebunden sind. Die Befunde lassen sich mit einer passiven, vor Interferenz schützenden Wirkung des Schlafes besser erklären.
- In einem zusätzlichen Methodenbeitrag wurde das Instrumentarium der kognitiven (Schlaf-) Forschung um eine Software erweitert, mit deren Hilfe sich äquivalente Lernmaterialien für die Anwendung von Messwiederholungsplänen vollautomatisch generieren lassen.

14. Literatur

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2, 396-408.

Baayen, R. H., Piepenbrock, R. & Gulikers, L. (1995). *The CELEX lexical database* (Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Barrett, T. R. & Ekstrand, B. R. (1972). Effect of sleep on memory III. Controlling for time-of-day effects. *Journal of Experimental Psychology*, 96, 321-327.

Benson, K. & Feinberg, I. (1975). Sleep and memory: retention 8 and 24 hours after initial learning. *Psychophysiology*, 12, 192-195.

Benson, K. & Feinberg, I. (1977). The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology*, 14, 375-384.

Blissit, P. A. (2001). Sleep, memory, and learning. *Journal of Neuroscience Nursing*, 33, 208-215.

Born, J., Kern, W., Bieber, K., Fehm-Wolfsdorf, G., Schiebe, M. & Fehm, H. L. (1986). Night-time plasma cortisol secretion is associated with specific sleep stages. *Biological Psychiatry*, 21, 1415-1424.

Brainerd, C. J. & Reyna, V. F. (2002). Fuzzy-Trace theory and false memory. *Current Directions in Psychological Science*, 1, 164-169.

Brainerd, C. J., Reyna, V. F. & Poole, D. A. (2000). Fuzzy-Trace theory and false memory: Memory theory in the courtroom. In D. F. Bjorklund (Ed.), *False memory creation in children and adults: Theory, research, and implications* (pp. 93-127). Mahwah, NJ: Erlbaum.

Bruce, D. & Winograd, E. (1998). Remembering Deese's 1959 articles: The Zeitgeist, the sociology of science, and false memories. *Psychonomic Bulletin & Review*, 5, 615-624.

Cipolli, C. (1995). Sleep, dreams and memory: An overview. *Journal of Sleep Research*, 4, 2-9.

Coenen, A. M. L. & Van Luitelaar, E. L. J. M. (1997). Effects of benzodiazepines, sleep and sleep deprivation on vigilance and memory. *Acta Neurologica Belgica*, 97, 123-129.

Crick, F. & Mitchison, G. (1983). The function of dream sleep. *Nature*, 304, 111-114.

Crick, F. & Mitchison, G. (1995). REM sleep and neural nets. *Behavioral Brain Research*, 69, 147-155.

Dahl, A. (1928). Über den Einfluss des Schlafens auf das Wiedererkennen. *Psychologische Forschung*, 11, 290-301.

Drosopoulos, S., Wagner, U. & Born, J. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory*, 12, 44-51.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50, 1096-1121.

Ebbinghaus, H. (1885, 1966). *Über das Gedächtnis*. Amsterdam: E. J. Bonset (Originalausgabe: Leipzig, Duncker & Humblot, 1885).

Ekstrand, B. R. (1967). Effect of sleep on memory. *Journal of Experimental Psychology*, 75, 64-72.

Ekstrand, B. R. (1972). To sleep, perchance to dream (about why we forget). In C. P. Duncan, L. Sechrest, A. W. Melton (Eds.), *Human memory. Festschrift in honor of Benton J Underwood* (pp. 59-82). New York: Appleton-Century-Crofts.

Ekstrand, B. R., Barrett, T. R., West, J. N. & Maier, W. G. (1977). The effect of sleep on human long-term memory. In R. R. Drucker-Colin, J. L. McGaugh (Eds.), *Neurobiology of sleep and memory* (pp. 419-438). New York: Academic Press.

Evans, C. R. & Newman, E. A. (1964). Dreaming: An analogy from computers. *New Scientist*, 419, 577-579.

Feinberg, I. & Evarts, E. (1969). Changing concepts of the function of sleep: Discovery of intense brain activity during sleep calls for revision of hypotheses as to its function. *Biological Psychiatry*, 1, 331-348.

Fenn, K. M., Nusbaum, H. C. & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature*, 425, 614-616.

Ficca, G., Lombardo, P., Rossi, L. & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioral Brain Research*, 112, 159-163.

Ficca, G. & Salzarulo, P. (2004). What in sleep is for memory. *Sleep Medicine*, 5, 225-230.

Fischer, S., Hallschmid, M., Elsner, A. L. & Born, J. (2002). Sleep forms memory for finger skills. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 11987-11991.

Gais, S., Plihal, W., Wagner, U. & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature Neuroscience*, 12, 1335-1339.

Graves, E. A. (1936). The effect of sleep upon retention. *Journal of Experimental Psychology*, 19, 316-322.

Hager, W. & Hasselhorn, M. (1994). *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.

Hinrichs, J. V., Ghoneim, M. M. & Mewaldt, S. P. (1984). Diazepam and memory: Retrograde facilitation produced by interference reduction. *Psychopharmacology*, 84: 158-162.

Hockey, G. R. J., Davies, S. & Gray, M. M. (1972). Forgetting as a function of sleep at different times of day. *Quarterly Journal of Experimental Psychology*, 24, 386-393.

Horne, J. A. (2000). REM sleep – by default? *Neuroscience and Biobehavioral Reviews*, 24, 777-797.

Howell, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury.

Huron, C., Servais, C & Danion, J. M. (2001). Lorazepam and diazepam impair true, but not false, recognition in healthy volunteers. *Psychopharmacology*, 155, 204-209.

Idzikowski, C. (1984). Sleep and memory. *British Journal of Psychology*, 75, 439-449.

Izura, C., Hernández-Muñoz, N. & Ellis, A. W. (2005). Category norms for 500 Spanish words in five semantic categories. *Behavior Research Methods*, 37, 385-397.

Jenkins, G. & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *American Journal of Psychology*, 35, 605-612.

Kirschbaum, C., Wolf, O. T., May, M., Wippich, W. & Hellhammer. D. H. (1996). Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy subjects. *Life Sciences*, 58, 1475-1483.

Krauth, J. (2000). *Experimental Design. A handbook and dictionary for medical and behavioral research*. Amsterdam: Elsevier.

Lahl, O. & Pietrowsky, R. (eingereicht a). Does the "Sleep Effect" on memory depend on sleep or on night time? *Sleep & Hypnosis*.

Lahl, O. & Pietrowsky, R. (eingereicht b). The long range effect of sleep on episodic memory. *Journal of Sleep Research*.

Lahl, O. & Pietrowsky, R. (eingereicht c). Effect of sleep on task-irrelevant stimuli. *Memory & Cognition*.

Lahl, O. & Pietrowsky, R. (im Druck). EQUIWORD: A software for the automatic creation of truly equivalent word lists. *Behavior Research Methods*.

Lahl, O., Zlomuzica, A. & Pietrowsky, R. (eingereicht). Sleep enhances memory for things that never happened. *Nature Neuroscience*.

Lazar, G. (1969). Retention as a function of successive recall trials following original learning. *Psychological Reports*, 25, 567-574.

Lazar, G. & Van Laer, J. (1966). Successive recall as a warm-up for paired adjectives. *Psychonomic Sciences*, 5, 137-138.

Lewin, I. L. & Glaubman, H. (1975). The effect of REM deprivation: Is it detrimental, beneficial, or neutral? *Psychophysiology*, 12, 349-353.

Lovatt, D. J. & Warr, P. B. (1968). Recall after sleep. *American Journal of Psychology*, 81, 253-257.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, 39, 352-370.

McGuigan, F. J. (1978). *Experimental Psychology. A methodological approach* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Meléndez, J., Galli, I., Boric, K., Orega, A., Zuñiga, L., Henríquez-Roklán C. F. et al. (2005). Zolpidem and triazolam do not affect the nocturnal sleep-induced memory improvement. *Psychopharmacology*, 181, 21-26.

Mueller, C. W., Lisman, S. A. & Spear, N. E. (1983). Alcohol enhancement of human memory: Tests of consolidation and interference hypotheses. *Psychopharmacology*, 80, 226-230.

Namboodiri, N. K. (1972). Experimental designs in which each subject is used repeatedly. *Psychological Bulletin*, 77, 54-64.

Nesca, M. & Koulack, D. (1994). Recognition memory, sleep and circadian rhythms. *Canadian Journal of Experimental Psychology*, 48, 359-379.

Newcomer, J. W., Craft, S., Hershey, T., Askins, K., Bardgett, M. E. (1994). Glucocorticoid-induced impairment in declarative memory performance in adult humans. *Journal of Neuroscience*, 14, 2047-2053.

Newman, E. B. (1939). Forgetting of meaningful material during sleep and waking. *American Journal of Psychology*, 52, 65-71.

Newman, E. A. & Evans, C. R. (1965). Human dream processes as analogous to computer programme clearance. *Nature*, 206, 534.

Nilsson, L. G., Law, J. & Tulving, E. (1988). Recognition failure of recallable unique names: Evidence for an empirical law of memory and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 266-277.

Osgood, C. E. & Suci, G. J. (1955). Factor analysis of meaning. *Journal of Experimental Psychology*, 50, 325-338.

Paivio, A., Yuille, J. C. & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, 76 (1, Pt. 2).

Parker, E. S., Birnbaum, I. M., Weingartner, H., Hartley, J. T., Stillman, R. C. & Wyatt, R. J. (1980). Retrograde enhancement of human memory with alcohol. *Psychopharmacology*, 69, 219-222.

Payne, D. G. (1986). Hypermnesia for pictures and words: Testing the recall level hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 16-29.

Payne, D. G., Elie, C. J., Blackwell, J. M. & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*, 35, 261-285.

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

Plihal, W. & Born, J. (1999a). Effects of early and late nocturnal sleep on priming and spatial memory. *Psychophysiology*, 36, 571-582.

Plihal, W. & Born, J. (1999b). memory consolidation in human sleep depends on inhibition of glucocorticoid release. *Neuroreport*, 10, 2741-2747.

Popper, K. R. (1968). *The logic of scientific discovery* (5th ed.). London: Hutchinson.

Portnoff, G., Baekeland, F., Goodenough, D. R., Karacan, I. & Shapiro, A. (1966). Retention of materials perceived immediately prior to onset of non-REM sleep. *Perceptual and Motor Skills*, 22, 751-758.

Rauchs, G., Desgranges, B., Foret, J. & Eustache, F. (2005). The relationship between memory systems and sleep stages. *Journal of Sleep Research*, 14, 123-140.

Richardson, A. & Gough, J. E. (1963). The long range effect of sleep on retention. *Australian Journal of Psychology*, 15, 37-41.

Richardson, J. & Gropper, M. S. (1964). Learning during recall trials. *Psychological Reports*, 15, 551-560.

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

Seamon, J. G., Luo, C. R., Kopecky, J. J., Price, C. A., Rothschild, L., Fung, N. S et al. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, 30, 1054-1064.

Siegel, J. M. (2001). The REM sleep-memory consolidation hypothesis. *Science*, 295, 1058-1063.

Spight, J. B. (1928). Day and night intervals and the distribution of practice. *Journal of Experimental Psychology*, 6, 397-398.

Stickgold, R., Whidbee, D., Schirmer, B., Patel, V. & Hobson, J. A. (2000). Visual discrimination task improvement: A multi-step process occurring during sleep. *Journal of Cognitive Neuroscience*, 12, 246-254.

Stickgold, R., Hobson, J. A., Fosse, R. & Fosse, M. (2001). Sleep, learning, and dreams: Off-line memory processing. *Science*, 294, 1052-1057.

Stickgold, R. & Walker, M. P. (2005). Sleep and memory: The ongoing debate. *Sleep*, 28, 1225-1227.

Toglia, M. P., Neuschatz, J. S. & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, 7, 233-256.

Tulving E. (1974). Cue-dependent forgetting. *American Scientist*, 62, 74-82.

Tulving, E. & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, 21, 437-484.

Tulving, E. & Psotka, J. (1971). Retroactive inhibition in free recall: Inaccessibility of information available in the memory store. *Journal of Experimental Psychology*, 87, 1-8.

Van Ormer, E. B. (1932). Retention after intervals of sleep and of waking. *Archives of Psychology*, 137, 1-49.

Vertes, R. P. (2004). Memory consolidation in sleep: Dream or reality. *Neuron*, 44, 134-148.

Vertes, R. P. & Siegel, J. M. (2005a). Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *Sleep*, 28, 1228-1229.

Vertes, R. P. & Siegel, J. M. (2005b). Rebuttal. *Sleep*, 28, 1232-1233.

Von Treuer, K., Norman, T. R. & Armstrong, S. M. (1996). Overnight human plasma melatonin, cortisol, prolactin, TSH, under conditions of normal sleep, sleep deprivation, and sleep recovery. *Journal of Pineal Research*, 20, 7-14.

Wagner, U., Gais, S. & Born, J. (2001). Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learning & Memory*, 8, 112-119.

Wagner, U., Gais, S., Haider, H., Verleger, R. & Born, J. (2004). Sleep inspires insight. *Nature*, 427, 352-355.

Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A. & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, 35, 205-211.

Walker, M. P. & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*, 44, 121-133.

Weitzman, E. D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T. F. & Hellman, L. (1971). Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *Journal of Clinical Endocrinology and Metabolism*, 33, 14-22.

Weitzman, E. D., Zimmerman, J. C., Czeisler, C. A. & Ronda, J. (1983). Cortisol secretion is inhibited during sleep in normal man. *Journal of Clinical Endocrinology and Metabolism*, 56, 352-358.

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, 20, 6-10.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, 55, 235-269.

Wolkowitz, O. M., Reus, V. I., Weingartner, H., Thompson, K., Breier, A., Doran, A. et al. (1990). Cognitive effects of corticosteroids. *American Journal of Psychiatry*, 147, 1297-1303.

Yaroush, R., Sullivan, M. J. & Ekstrand, B. R. (1971). Effects of sleep on memory II: Differential effect of the first and second half of the night. *Journal of Experimental Psychology*, 88, 361-366.

Zacks, R. T. & Hasher, L. (1994). Directed ignoring: Inhibition regulation of working memory. In D. Dagenbach, T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 241-264). San Diego: Academic Press.

Does the "Sleep Effect" on Memory depend on Sleep or on Night Time?

Olaf Lahl and Reinhard Pietrowsky

Institute of Experimental Psychology, University of Düsseldorf, Germany

Corresponding Author:

Olaf Lahl
Universität Düsseldorf
Klinische Psychologie
Universitätsstr. 1
40225 Düsseldorf
Germany
Phone: ++49-211-811-2146
Fax: ++49-211-811-4261
E-mail: olaf.lahl@uni-duesseldorf.de

Abstract

Numerous investigations demonstrated superior verbal memory performance after retention intervals of nocturnal sleep as opposed to diurnal wakefulness. However, it is not clear if the effect is attributable to either sleep or circadian phase. The present study therefore examined verbal recall after retention intervals of nocturnal sleep, diurnal wakefulness, and nocturnal wakefulness (sleep deprivation). Forty university students (range 19-30 years) were randomly assigned to one of the three conditions and were tested for cued recall of a paired-associate list 7 h after original learning. In line with previous findings, subjects in the nocturnal sleep condition expressed superior recall when compared to subjects in the diurnal wakefulness condition. However, contrary to predictions, recall performance between the nocturnal sleep and the nocturnal wakefulness condition did not differ significantly. The results raise some doubt on the generalizability of the beneficial effect of sleep on memory.

Key words: sleep, sleep deprivation, memory, retention, circadian phase

Introduction

It is a well-established fact that recall of verbal information from long-term memory is superior after retention intervals of nocturnal sleep in comparison to retention intervals of daytime wakefulness (1-9). This finding was subsequently termed “sleep effect” (10). However, due to the confounding of the two factors sleep/wake and circadian phase that is inherent in the experimental design used, the effect cannot be attributed unambiguously to the independent variable sleep. This leaves the question whether the beneficial effect on memory is mediated by sleep per se or whether it “may actually be due to circadian variables which simply share the same period of time” (6, p. 372). To disentangle both factors, the sleep and the wake condition have to occupy the same nocturnal period.

Though primarily aimed at isolating the differential effects of slow wave sleep (SWS) and rapid eye movement (REM) sleep, partial evidence for a beneficial effect of sleep per se comes from research investigating the first and second half of a 7 h nocturnal sleep separately. With the exception of Wagner et al. (11) all studies in this area found significantly better retention when subjects slept through the first 3-4 h of the night than when they were kept awake during the same nocturnal period (10, 12-17). However, from these investigations no conclusion about the role of the naturally occurring entire night sleep cycle can be drawn.

Hockey et al. (18) conducted an experiment with four different 5 h retention intervals. Two groups learnt a list of nouns in the late evening and recalled it in the late night. Two other groups had their learning and recall in the early and late morning respectively. One group of each time of day condition was allowed to sleep while the others underwent sleep deprivation during the retention interval. The results indicated a main effect for time of day but not for sleep, thus favoring circadian factors rather than sleep as the critical variable.

To our knowledge, only two studies attempted to compare a full night of sleep with a corresponding period of total sleep deprivation between learning and recall. Idzikowski (19) had subjects learn a list of nonsense syllables in the morning. After 16 h of subsequent daily activity one group went to sleep for 8 h while the other group underwent total sleep deprivation during the same period of night. Despite 16 h of waking activity that preceded the critical night period, free recall, paced recall, and relearning in the morning showed significantly better results for the sleep group. Nesca and Koulack (6) on the other hand did not find superior memory for the sleep condition on a verbal recognition task when comparing 8 h nocturnal retention intervals that were either filled completely with sleep or wakefulness.

Another critical issue in the research of sleep and memory that has widely been neglected is the problem of insufficient measures of retrieval from long-term memory. As pointed out by Roediger and Guynn “a single test of memory is an imperfect indicator of knowledge” (20, p. 200). Contrary to the well-known forgetting curves of the pioneering work of Ebbinghaus (21), which indicate a monotonous decline of memory with the passage of time, numerous studies have demonstrated an *increase* of recall between consecutive tests (22-26). Although the effect of sleep on memory has been investigated by various methods of retrieval (relearning, free recall, cued recall, recognition), only Barrett and Ekstrand (12) adopted the method of repeated recall tests. When comparing retention after sleep during the first and second half of the night, these authors found significantly better recall after the first half on the first test, but not on the second. These results demonstrate that the analysis of only one single recall test is likely to underestimate the amount of information stored in memory and may thus yield misleading results. A similar reason for possibly inaccurate measures of retention is the failure to control for report bias. While it is common practice in recognition tasks to have one measure for the ability to discriminate old and new items and one for a potential response bias (27, 28), standard cued and free recall paradigms usually do not include a separate measure of report bias. Thus, standard recall procedures probably underestimate retention of those subjects that have a conservative threshold and therefore tend to withhold a response unless they are quite sure it is the correct one. Such underestimation can easily be circumvented by forced report techniques, which force or at least encourage subjects to guess a response rather than to give no response at all in case of uncertainty (29, 30).

The aim of the present study was to further clarify the possible constraints of the beneficial effect of sleep on memory under conditions of circadian phasing and appropriate measures of memory retrieval. We therefore compared memory performance over three different retention intervals of nocturnal sleep, diurnal waking, and nocturnal waking. To effectively exhaust storage in long-term memory, we adopted successive recall trials and, beginning with the second trial, asked subjects to guess if they were not sure about the correct response. Based on the existing literature we expected to replicate the classical sleep memory effect, that is better retention after nocturnal sleep than after diurnal wakefulness. The critical question was however, whether sleep would still lead to superior recall with circadian factors held constant, i.e. better retention after nocturnal sleep than after nocturnal wakefulness.

Method

Subjects

Forty university students (19 male, 21 female) aged between 19-30 years (M = 24.3, SD = 2.9) participated in the experiment for financial compensation and were randomly assigned to one of the three experimental groups Sleep/Night (n = 12), Wake/Night (n = 12), and Wake/Day (n =

16). Inclusion criteria were non-smoking, regular sleep schedule (7-8 hours nocturnal sleep, sleep latency below 30 min, no excessive daytime napping), absence of any psychoactive medication and sleep disturbances during the last four weeks, no history of neurological or psychiatric illness. Subjects were obliged to refrain from alcoholic beverages, caffeine, and daytime napping 12 h before conducting the experiment. All participants gave signed consent to take part in the study after the study protocol had been fully explained.

Memory testing

For the sake of comparability with a great deal of previous research, memory was probed by a paired-associate list (PAL) of 16 noun pairs. To keep guessing probability low, the stimulus and response words of all pairs were weak associates. All words contained 5-9 letters and had a moderate imagery rating (*z*-scores between 0.00 and 0.53) (31). The complete list of words is presented in the Appendix.

Learning was carried out by the study-test method in four consecutive trials. During study trials a computer program presented all 16 pairs in sequential random order with a presentation rate of 1/1500 ms and an interstimulus interval of 1000 ms. During test trials the stimulus words only were presented in sequential random order and subjects had to type the matching response word on a keyboard. Except for the last test trial, which served as a measure of original learning, all responses were followed by feedback about their correctness (displaying “Correct” or “False” for 1 sec). To prevent active rehearsal, subjects had to perform simple additions of three digit numbers for 2 min before and after the last test trial.

The recall session consisted of four consecutive test trials without feedback. Beginning with the second recall test trial subjects were instructed to guess the correct response if they were unsure about it. Memory testing was completed by a final forced-list recognition test (32) that contained all 16 response words together with 16 matching lures of semantic similarity (see Appendix). The computer program simultaneously presented all response words and lures in random order on the left side of the screen. The subjects' task was then to use the computer mouse to move those 16 words to the panel on the right side that they felt were the response words presented during learning.

Design and procedure

The length of the retention interval was 7 h for all three groups. This period was chosen because recent epidemiological surveys (33-35) and laboratory findings (36) demonstrated that seven hours is the average habitual sleep period of healthy young adults.

To ensure that all subjects were fully awake during test sessions, they had to perform two computerized psychomotor tracking tasks of 3 min duration each, before learning and recall began (see (37) for a similar attempt using a mirror tracing task). The first task required them to follow random movements of a marker on the screen with the computer mouse as closely as possible. On the second task, subjects saw a vertical bar between two fixed bars on the left and right side of the screen. Due to unpredictable "forces" simulated by the software, the bar continuously tended to move away from the center location in one or the other direction. The task was to compensate those disturbances by moving the computer mouse in the opposite direction thereby avoiding the middle bar hitting one of the outer bars. Each such hit was signaled by a short beep and was scored as an error. However, since fully awakening was the sole reason to deploy both tracking tasks, no data regarding them will be reported in the Results section.

Sleepiness was measured subsequently during learning and recall sessions by a 60 sec finger-tapping task. This task is quickly and easily accomplished by subjects. Yet, it has shown sensitivity to fatigue as indicated by a decline in tapping rate due to hangover effects of sedative drugs (38-41). Subjects had to strike the Enter key of the numerical keyboard block repetitively for 60 sec as quickly as possible. In order to avoid premature disengagement, they were in fact misinformed that they had to tap for 90 sec. However, the software terminated processing keyboard input already after 60 sec. In addition, subjective sleepiness was assessed by the Stanford Sleepiness Scale (SSS) (42, 43).

Subjects in the sleep condition were acclimated to the placement of electrodes and the sleep laboratory by spending one adaptation night. To ensure comparability with the two other conditions, the adaptation night was always scheduled two nights before the test night, so that subjects of all groups would sleep at home the night right before the experiment. In the experimental night, they reported to the laboratory at 22.00 h for placement of electrodes and began memory testing between 23.15 h and 23.30 h, which took about 30 min. Lights were turned off between 23.45 h and 00.00 h, immediately after original learning was completed. Sleep was monitored all night by polysomnography according to standard criteria (44). Awakening from sleep was between 06.45 h and 07.00 h. The retest session began between 07.00 h and 07.15 h.

Subjects in the Wake/Night condition performed learning and recall at 23.30 h and 07.00 h respectively. During the 7 h retention interval, they stayed awake under the control of always two experimenters. They were free to watch videos, read recreational materials, or talk with the experimenters. Consumption of caffeine or alcoholic beverages was prohibited throughout the entire night. Subjects in the daytime condition began learning at 08.30 h. Following this, they were obliged to refrain from caffeine, alcohol, and napping and were dismissed from the laboratory. They returned for recall at 16.00 h.

To prevent any exchange of learning material, all subjects participated individually. Subjects who failed to recall at least four items of the paired associate list during the last trial of original learning were excluded from further analysis and replaced by others. Likewise, subjects in the sleep group who exhibited a sleep latency above 45 min or a sleep efficiency index below 80% were substituted by others. Three participants in the Sleep/Night and Wake/Night group respectively were replaced because they had failed to meet the learning criterion. Two participants in the Sleep/Night group were replaced due to poor sleep quality.

Data analysis

Sleep recordings were analyzed off-line according to standard criteria (44). Relevant sleep parameters were sleep onset latency, amounts of sleep stages 1, 2, SWS, and REM in relation to time spent in bed, and sleep efficiency index (time asleep per time in bed). Original learning was defined as the number of correct items on the last learning trial. Two measures of recall were assessed. The number of items given in the first recall trial adhered to the common single recall score. Extended recall on the other hand was measured by the cumulative count of correct responses over all of the four recall trials. That is, a correct response was scored once the first time it was produced while ignoring subsequent repetitions. Forgetting was then defined as the difference between a recall score and the original learning score divided by the original learning score. Multiplication by 100 yielded the percentage of items lost over the retention interval. Recognition performance was assessed by the number of correct choices (hits). Due to the application of forced-list recognition, no extra measure of response bias was necessary (45).

Pairwise Dunnett *t* tests (46), denoted t_d, for multiple comparisons with a single control (i.e. the sleep condition) were calculated for finger tapping rates, sleepiness ratings (two-tailed tests), percent loss scores, and recognition hits (one-tailed tests) to compare sleepiness and memory performance between the Wake/Day and Sleep/Night and the Wake/Night and

Sleep/Night conditions respectively. Unless specified otherwise, all data presented in the Results section are expressed as means ± standard error of mean (*SEM*).

Results

Sleep parameters and sleepiness

Sleep parameters in the experimental night (Table 1) were within the normal range of healthy young subjects (47-50) indicating successful realization of the sleep condition.

Insert Table 1 about here

Finger tapping rates (Figure 1A) did not indicate substantial differences between conditions at learning (Wake/Day: 302.88 ± 9.85; Sleep/Night: 300.00 ± 17.44; Wake/Night: 316.50 ± 13.18; $t_d(37) = 0.16$ for comparison Wake/Day vs. Sleep/Night and $t_d(37) = 0.83$ for comparison Wake/Night vs. Sleep/Night) and recall (Wake/Day: 306.31 ± 11.50; Sleep/Night: 302.58 ± 14.41; Wake/Night: 294.67 ± 13.71; $t_d(37) = 0.20$ for comparison Wake/Day vs. Sleep/Night; $t_d(37) = 0.41$ for comparison Wake/Night vs. Sleep/Night). Subjective sleepiness ratings (Figure 1B) were comparable at learning (Wake/Day: 2.00 ± 0.20; Sleep/Night: 2.50 ± 0.31; Wake/Night: 2.17 ± 0.27; $t_d(37) = 1.39$; $p = .14$ for comparison Wake/Day vs. Sleep/Night; $t_d(22) = 0.87$ for comparison Wake/Night vs. Sleep/Night) but showed considerable differences at recall (Wake/Day: 1.88 ± 0.24; Sleep/Night: 3.25 ± 0.28; Wake/Night: 4.75 ± 0.35; $t_d(37) =$

3.45; p = .002 for comparison Wake/Day vs. Sleep/Night; t_d (37) = 3.52; p = .001 for comparison Wake/Night vs. Sleep/Night).

Insert Figure 1 about here

Memory measures

Figure 2 illustrates the number of items recalled in the paired-associate task at three different stages of memory testing: (i) the last test trial during learning, which marks the original learning score (OL), (ii) the first test trial during recall (T1), and (iii) the last test trial during recall, which is expressed as cumulative score over all test trials (T1-T4 cum.). Levels of original learning were 9.69 ± 0.73 for the Wake/Day, 10.67 ± 1.07 for the Wake/Night, and 10.00 ± 0.83 for the Sleep/Night condition. Between the last learning trial and the first recall trial, memory performance declined markedly in the Wake/Day condition (2.12 ± 0.54 items lost) and to a lesser extent in the two night conditions (Wake/Night: 0.92 ± 0.26; Sleep/Night: 0.83 ± 0.39 items lost). However, forgetting was alleviated during successive recall in all groups, which led to a recovery of 9.97% ± 2.68% (Wake/Day), 8.19% ± 4.13% (Wake/Night), and 8.09% ± 3.66% (Sleep/Night) of the items originally not recalled on T1.

Insert Figure 2 about here

Proportions of items lost over the retention interval are shown in Figure 3 for T1 and T1-T4 cum. The main finding holds for both measures: subjects in the Wake/Day condition lost substantially more items (T1: 25.89% ± 7.21%; T1-T4 cum.: 15.92 % ± 6.86%) than did subjects in the Sleep/Night condition (T1: 7.58% ± 4.19%; T1-T4 cum.: –0.52% ± 3.82%) yielding test statistics of $t_d(37) = 2.19$ ($p = .031$) for T1 and t_d $(37) = 2.09$ ($p = .039$) for T1-T4 cum. However, subjects in the Wake/Night condition (T1: 11.98% ± 4.67%; T1-T4 cum.: 3.79% ± 4.35%) did not loose significantly more items than those in the Sleep/Night condition (T1: t_d $(37) = 0.49$; T1-T4 cum.: t_d $(37) = 0.51$).

Insert Figure 3 about here

The number of recognition hits was generally high under all conditions (Wake/Day: 13.31 ± 0.60; Wake/Night: 14.08 ± 0.40; Sleep/Night: 14.00 ± 0.30). There were no substantial differences between groups for this measure (t_d $(26) = 1{,}01$; $p = .25$ for comparison Wake/Day vs. Sleep/Night, $t(22) = 0.11$; $p = 0.70$ for comparison Wake/Night vs. Sleep/Night).

Discussion

In accordance with previous research, the results of the present study confirmed superior declarative retention after a period of night sleep in comparison to a period of daytime activity. Subjects in the Wake/Day condition lost one-fourth of list associations on the first recall attempt and still 16% after cumulating over all recall trials while those in the Sleep/Night condition in fact showed a minimal increase (0.52%) of items or *negative forgetting* in the cumulative count. Based on the assumption that this effect is due to sleep per se, it was predicted that night sleep would also lead to better memory performance when compared to a condition of nightly waking activity. However, our data do not corroborate this hypothesis. Retention did not differ significantly between the two night conditions. This result is even more remarkable in view of the sleepiness data we obtained. While ratings of subjective sleepiness did not differ at learning, subjects in the Wake/Night condition exhibited the expected dramatic rise in sleepiness over night. On the average, they felt “foggy, losing interest in remaining awake” whereas subjects in the Sleep/Night condition were “awake, but relaxed; responsive but not fully alert” and those in the Wake/Day condition were “functioning at high level, but not at peak; able to concentrate”. Thus, despite the fact that subjects in the sleep deprivation group acted at levels of severe drowsiness, they exhibited recall levels comparable to those in the sleep condition.

The application of successive recall trials in conjunction with guessing instructions from the second recall trial led to an improvement of memory performance in all groups. To what extent this partial recovery of items is related to one or the other factor cannot be determined and was not an aim of the present study. While we replicated the important finding, that a single recall trial is an insufficient measure of memory, there was no indication of a differential benefit between groups due to repeated recall. All groups recovered about 8-10% of the items so that the ranking of group performances was not affected by the extended recall procedure. Recognition performance did not differ significantly between groups albeit the ranking of group performances was the same as for recall.

With the present investigation we attempted to assess the effect of sleep on memory under various methodological precautions including application of paired-associate learning, which is known to be sensitive for the effect of sleep; exhaustion of long-term storage by extended recall; Wake/Day control group to replicate previous results; polysomnographic control of night sleep; separate accomplishment of sleep deprivation to prevent exchange of learning material between subjects. Notwithstanding these strict methodologies, our results were negative with regard to a sleep memory effect under conditions of circadian phasing, as were those obtained by Nesca and Koulack (6). It should be noted that the effect of sleep on retention of pair associations is usually strong enough to prove statistically valid with 10-16 subjects per condition (3), even when the retention interval covers only the first half of the night (12, 13, 15, 17). Thus, although generally no direct conclusion can be drawn from a null result, our findings may put some constraints on the generalizability of the beneficial effect of sleep on retention.

The comparable retention levels under both night conditions cannot easily be reconciled with the results obtained by Idzikowski (19) who found a large effect on retention of sleep per se. It is however in line with the findings of Hockey et al. (18) and Nesca & Koulack (6). An apparent explanation for the non-difference between the two night conditions in these studies and in ours could be that circadian factors rather than sleep play a major role in controlling what is retained in long-term memory and what is not. Above all, the glucocorticoid cortisol seems the best candidate to mediate circadian fluctuations of memory performance. It has a well-established detrimental effect on the declarative (hippocampus-mediated) memory system (51-55) while its secretion follows a pronounced circadian rhythmicity with its nadir in the first and its peak in the second half of the night (56, 57). However, our results cannot readily be explained by the action of cortisol since experimental evidence indicates an increase of cortisol levels under conditions of sleep deprivation as opposed to nocturnal sleep (58, 59). This increase should have led to a reduction of memory performance in the Wake/Night group, which we did not observe. The comparable levels of retention under both night conditions would be best explained by a factor that solely depends on circadian timing, but not on sleep. To our knowledge, empirical evidence for such a factor is yet outstanding.

References

1. Benson K, Feinberg I. Sleep and memory: Retention 8 and 24 hours after initial learning. Psychophysiology 1975;12:192-195.

2. Benson K, Feinberg I. The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. Psychophysiology 1977;14:375-384.

3. Ekstrand BR. Effect of sleep on memory. J Exp Psychol 1967;75:64-72.

4. Jenkins G, Dallenbach KM. Obliviscence during sleep and waking. Am J Psychol 1924;35:605-612.

5. Lovatt DJ, Warr PB. Recall after sleep. Am J Psychol 1968;81:253-257.

6. Nesca M, Koulack D. Recognition memory, sleep and circadian rhythms. Can J Exp Psychol 1994;48:359-379.

7. Newman EB. Forgetting of meaningful material during sleep and waking. Am J Psychol 1939;52:65-71.

8. Spight JB. Day and night intervals and the distribution of practice. J Exp Psychol 1928;6:397-398.

9. Van Ormer EB. Retention after intervals of sleep and of waking. Arch Psychol 1932;137:1-49.

10. Ekstrand BR, Barrett TR, West JN, Maier WG. The effect of sleep on human long-term memory. In: Drucker-Colin RR, McGaugh JL, eds. Neurobiology of sleep and memory. New York: Academic Press, 1977; 419-438.

11. Wagner U, Gais S, Born J. Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. Learn Mem 2001;8:112-119.

12. Barrett TR, Ekstrand BR. Effect of sleep on memory III. Controlling for time-of-day effects. J Exp Psychol 1972;96:321-327.

13. Fowler MJ, Sullivan MJ, Ekstrand BR. Sleep and memory. Science 1973;179:302-304.

14. Gais S, Born J. Low acetylcholine during slow wave sleep is critical for declarative memory consolidation. P Natl Acad Sci USA 2004;101:2140-2144.

15. Grosvenor A, Lack LC. The effect of sleep before or after learning on memory. Sleep 1984;7:155-167.

16. Plihal W, Born J. Effects of early and late nocturnal sleep on declarative and procedural memory. J Cogn Neurosci 1997;9:534-547.

17. Yaroush R, Sullivan MJ, Ekstrand BR. Effects of sleep on memory II: differential effect of the first and second half of the night. J Exp Psychol 1971;88:361-366.

18. Hockey GRJ, Davies S, Gray MM. Forgetting as a function of sleep at different times of day. Q J Exp Psychol 1972;24:386-393.

19. Idzikowski C. Sleep and memory. Br J Psychol 1984;75:439-449.

20. Roediger HL, Guynn ML. Retrieval processes. In: Bjork EL, Bjork RA, eds. Memory. San Diego: Academic Press, 1996;197-236.

21. Ebbinghaus H. Über das Gedächtnis, Leipzig: Duncker & Humblot, 1885.

22. Brown W. To what extent is memory measured by a single recall? J Exp Psychol 1923;6:377-382.

23. Lazar G. Retention as a function of successive recall trials following original learning. Psychol Rep 1969;25:567-574.

24. Lazar G., Van Laer J. Successive recall as a warm-up task for paired adjectives. Psychon Sci 1966;5:137-138.

25. Payne DG. Hypermnesia for pictures and words: Testing the recall level hypothesis. J Exp Psychol Learn Mem Cogn 1986;12:16-29.

26. Richardson J, Gropper MS. Learning during recall trials. Psychol Rep 1964; 15:551-560.

27. Neath I, Surprenant AM. Human memory: An introduction to research, data, and theory, Belmont, CA: Wadsworth, 2003.

28. Snodgrass JG, Corwin J. Pragmatics of measuring recognition memory: Applications to dementia and amnesia. J Exp Psychol Gen 1988;117:34-50.

29. Erdelyi MH, Finks J, Feigin-Pfau MB. The effect of response bias on recall performance, with some observations on processing bias. J Exp Psychol Gen 1989;118:245-254.

30. Higham PA. Strong cues are not necessariliy weak: Thomson & Tulving (1970) and the encoding specifity principle revisited. Mem Cognit 2002;30:67-80.

31. Baschek IL, Bredenkamp J, Oehrle B, Wippich W. Bestimmung der Bildhaftigkeit, Konkretheit und der Bedeutungshaltigkeit von 800 Substantiven. In: Hager W, Hasselhorn M, eds. Handbuch deutschsprachiger Wortnormen. Göttingen: Hogrefe, 1994;174-186.

32. Brown J. An analysis of recognition and recall and of problems in their comparison. In: Brown J, ed. Recall and recognition. London: John Wiley & Sons Ltd, 1976;1-35.

33. Ban DJ, Lee TJ. Sleep duration, subjective sleep disturbances and associated factors among university students in Korea. J Korean Med Sci 2001;16:475-480.

34. Breslau N, Roth T, Rosenthal MD, Andreski MA. Daytime sleepiness: An epidemiological study of young adults. Am J Public Health 1997;87:1649-1653.

35. Hicks RA, Lucero-Gorman K, Bautista J, Hicks GJ. Ethnicity, sleep duration, and sleep satisfaction. Percept Mot Skills 1999;88:234-235.

36. Hirshkowitz M, Moore CA, Hamilton CR, Rando KC, Karacan I. Polysomnography of adults and elderly: Sleep architecture, respiration, and leg movement. J Clin Neurophysiol 1992;9:56-62.

37. Yaroush R, Sullivan MJ, Ekstrand BR. Effects of sleep on memory II: Differential effect of the first and second half of the night. J Exp Psychol 1971;88:361-366.

38. Bond AJ, Lader MH. Residual effects of hypnotics. Psychopharmacologia 1972; 25:117-132.

39. Kornetsky C, Vates TS, Kessler EK. A comparison of hypnotic and residual psychological effects of single doses of chlorpromazine and secobarbital in man. J Pharmacol Exp Ther 1959;127:51-54.

40. Peck AW, Adams R, Bye C, Wilkinson RT. Residual effects of hypnotic drugs: evidence for individual differences on vigilance. Psychopharmacology 1976;47:213-216.

41. Walters AJ, Lader MH. Hangover effects of hypnotics on man. Nature 1971;229:637-638.

42. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: A new approach. Psychophysiology 1973;10:431-436.

43. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Stanford Schläfrigkeitsskala (Stanford Sleepiness Scale) Selbstbeurteilungs-Skala. In: Collegium Internationale Psychiatriae Scalarum, eds. Internationale Skalen für Psychiatrie, Göttingen: Beltz-Test, 1996;137-139.

44. Rechtschaffen A, Kales A. A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects, Washington: Public Health Service Publications, 1968.

45. Brown J, Packham DW. The effect of prior recall on multiple response recognition. Q J Exp Psychol 1967;19:356-361.

46. Dunnett CW. A multiple comparison procedure for comparing several treatments with a control. J Am Stat Assoc 1955;50:1096-1121.

47. Browman CP, Cartwright CD. The first-night effect on sleep and dreams. Biol Psychiatry 1980;15:809-812.

48. Carskadon MA, Dement WC. Normal human sleep: An overview. In: Kryger MH, Roth T, Dement WC, eds. Principles and practice of sleep medicine, Philadelphia: W. B. Saunders Company, 1989;3-13.

49. Clausen J, Sersen EA, Lidsky A. Variability of sleep measures in normal subjects. Psychophysiology 1974;11:509-516.

50. Lorenzo JL, Barbanoj MJ. Variability of sleep parameters across multiple laboratory sessions in healthy young subjects: The “very first night effect”. Psychophysiology 2002;39:409-413.

51. Kirschbaum C, Wolf OT, May M, Wippich W, Hellhammer DH. Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy subjects. Life Sci 1996;58:1475-1483.

52. Newcomer JW, Craft S, Hershey T, Askins K, Bardgett ME. Glucocorticoid-induced impairment in declarative memory performance in adult humans. J Neurosci 1994;14:2047-2053.

53. Plihal W, Born J. Memory consolidation in human sleep depends on inhibition of glucocorticoid Release. Neuroreport 1999;10:2741-2747.

54. Plihal W, Pietrowsky R, Born J. Dexamethasone blocks sleep induced improvement of declarative memory. Psychoneuroendocrinology 1999;24:313-331.

55. Wolkowitz OM, Reus VI, Weingartner H et al. Cognitive effects of corticosteroids. Am J Psychiatry 1990;147:1297-1303.

56. Born J, Kern W, Bieber K. Night-time plasma cortisol secretion is associated with specific sleep stages. Biol Psychiatry 1986;21:1415-1424.

57. Weitzman ED, Fukushima D, Nogeire C, Roffwarg H, Gallagher TF, Hellman L. Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. J Clin Endocrinol Metab 1971;33:14-22.

58. Von Treuer K, Norman TR, Armstrong SM. Overnight human plasma melatonin, cortisol, prolactin, TSH, under conditions of normal sleep, sleep deprivation, and sleep recovery. J Pineal Res 1996;20:7-14.

59. Weitzman ED, Zimmerman JC, Czeisler CA, Ronda J. Cortisol secretion is inhibited during sleep in normal man. J Clin Endocrinol Metab 1983;56:352-358.

Tables

Table 1: Sleep parameters in the experimental night

Sleep parameter	*Mean*	*SEM*
Time in bed [min]	414.96	3.13
S2-Sleep onset latency [min]	19.17	2.98
Sleep efficiency	92.09	1.36
% Wake	7.89	1.36
% S1	9.28	1.16
% S2	41.21	1.46
% SWS	20.72	1.33
% REM	17.79	1.41
% MT	3.10	0.78

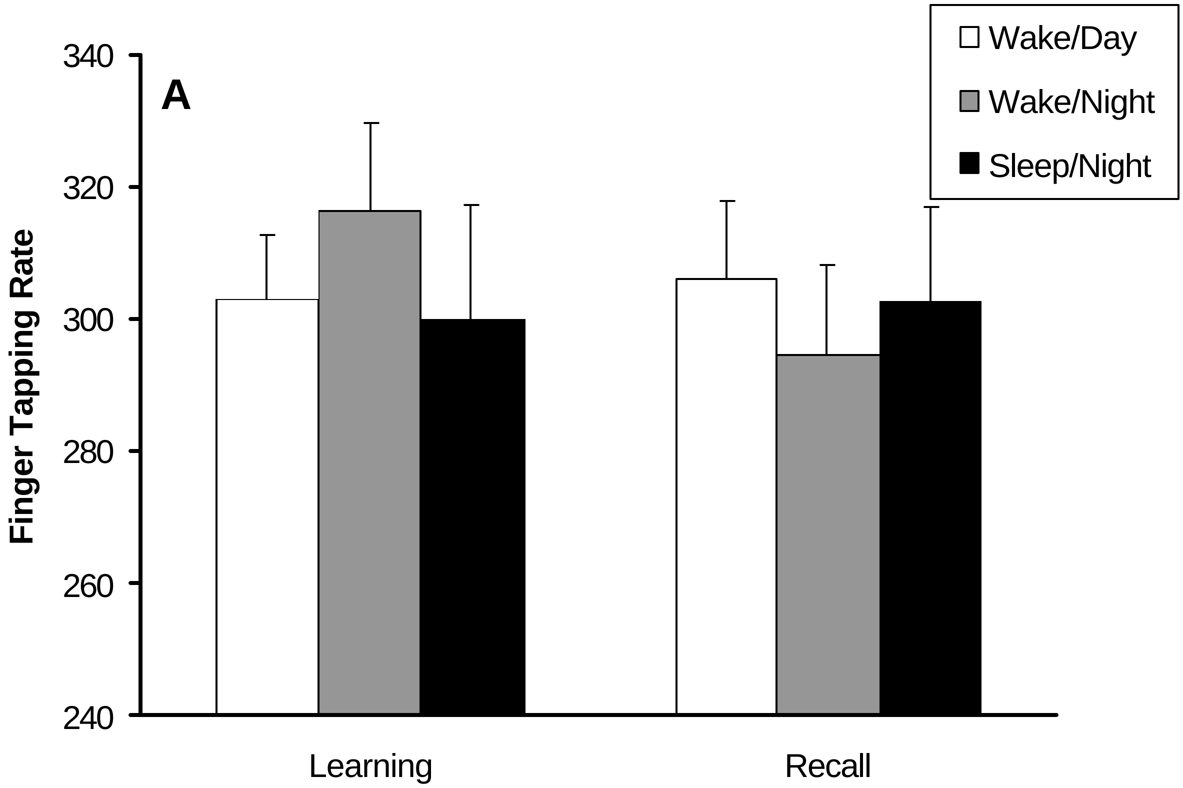
Note. S1: sleep stage 1; S2: sleep stage 2; SWS: slow wave sleep; REM: rapid eye movement sleep; MT: movement time. Percentages are relative to time in bed.

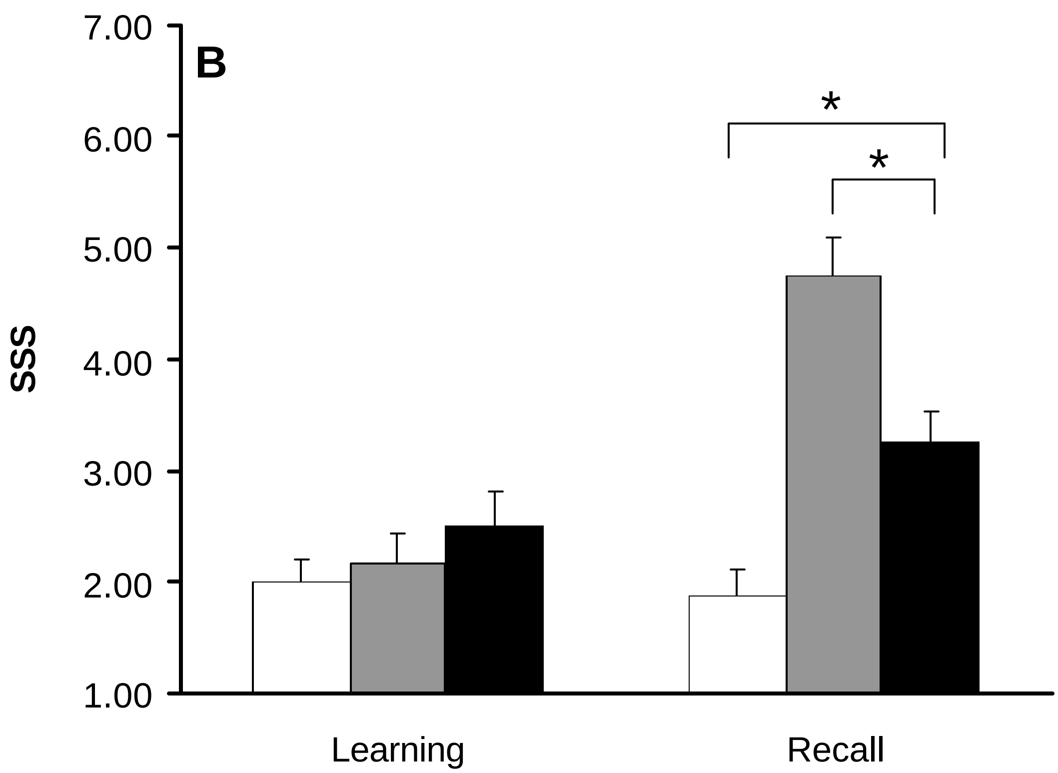
Figure Captions

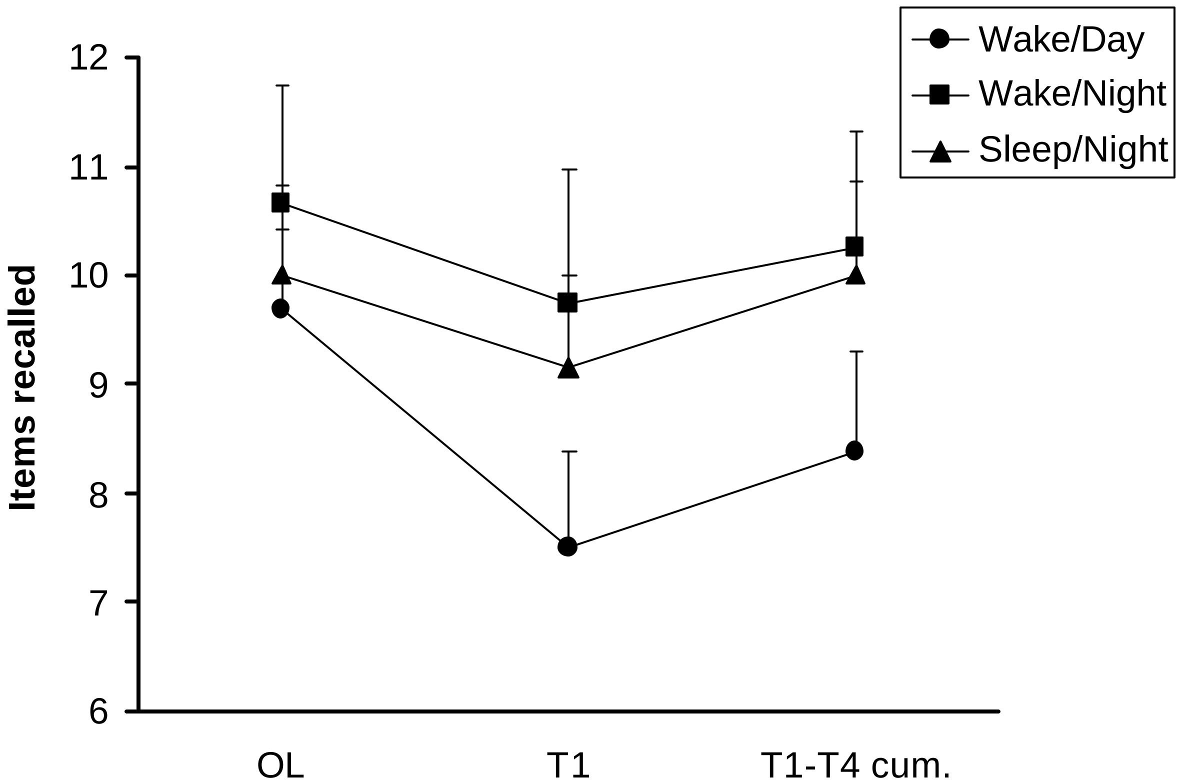
Figure 1. Control variables at learning and recall for the three experimental conditions. **A** Finger tapping rates (*M* ± *SEM*). **B** Ratings on the Stanford Sleepiness Scale (medians).

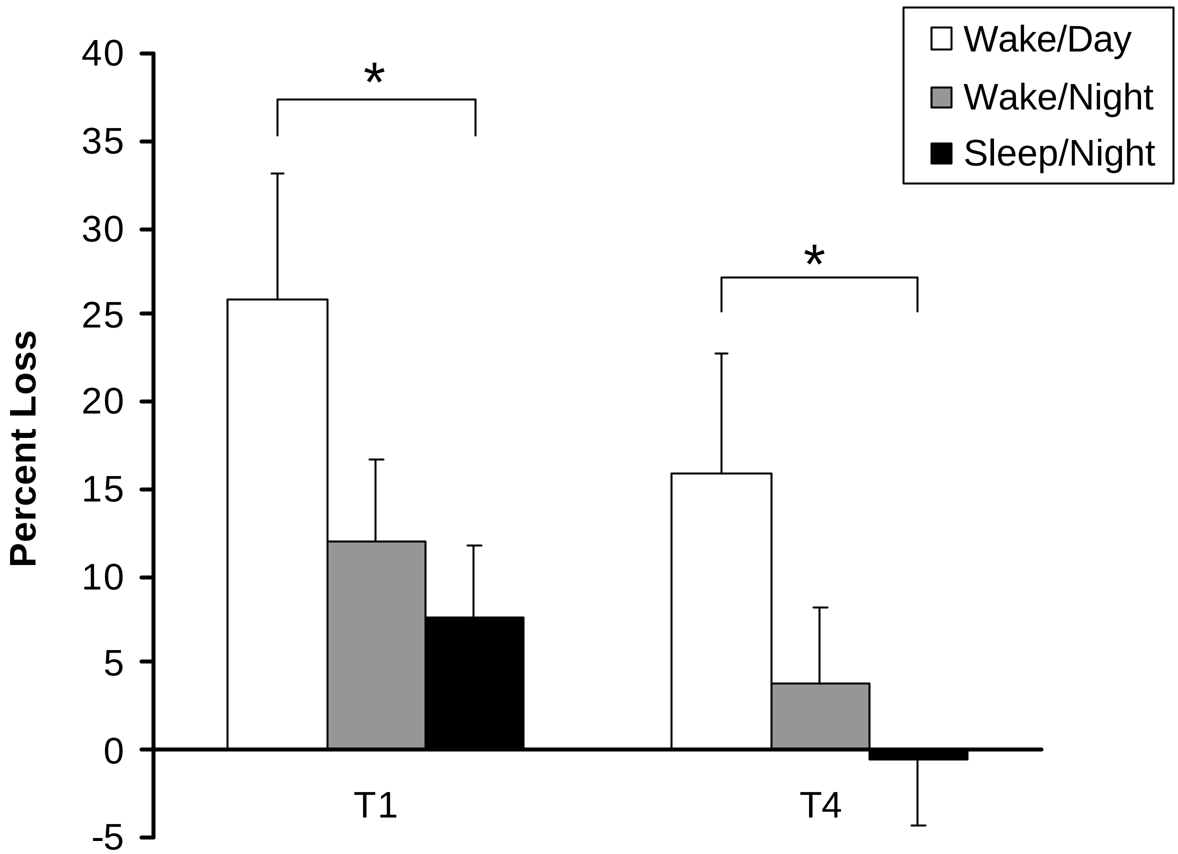
Figure 2. Number of items recalled on the last learning trial (OL), the first recall trial after the 7 h retention interval (T1), and cumulatively across all four consecutive recall trials (T1-T4 cum.). Error bars indicate *SEM*.

Figure 3. Percentage of items lost between original learning and the first recall trial (T1) and between original learning and the last recall trial (T1-T4 cum.). Scores at T1-T4 cum. are cumulative across all four consecutive recall trials. Error bars indicate *SEM*.


Wake/Day
Wake/Night
Sleep/Night
340
320
300
280
260
240
A
Finger Tapping Rate
Learning
Recall


7.00
6.00
5.00
4.00
3.00
2.00
1.00
B
*
*
SSS
Learning
Recall


Wake/Day
Wake/Night
Sleep/Night
12
11
10
9
8
7
6
Items recalled
OL
T1
T1-T4 cum.


Wake/Day
Wake/Night
Sleep/Night
Percent Loss
40
35
30
25
20
15
10
5
0
-5
*
*
T1
T4

Appendix: Paired-associate list and recognition lures for response words (English translation in parentheses)

Stimulus	Response	Distractor
Beweis (proof)	Wirkung (effect)	Ergebnis (result)
Verein (club)	Beginn (beginning)	Anfang (origin)
Stimmung (mood)	Auswahl (choice)	Teilmenge (subset)
Geschöpf (creature)	Interesse (interest)	Gefallen (favour)
Antwort (answer)	Vertreter (agent)	Ersatzmann (substitute)
Haltung (attitude)	Beitrag (contribution)	Mitwirkung (participation)
Ferne (distance)	Inhalt (content)	Begriff (idea)
Aufgabe (task)	Richtung (direction)	Ziel (target)
Kosten (costs)	Geruch (smell)	Duft (fragrance)
Neffe (nephew)	Besitz (possession)	Eigentum (property)
Versuch (attempt)	Gebet (prayer)	Andacht (devotion)
Beruf (profession)	Monat (month)	Jahr (year)
Merkmal (characteristic)	Vortrag (lecture)	Rede (speech)
Bewohner (resident)	Kreislauf (circulation)	Umdrehung (rotation)
Profil (profile)	Gruppe (group)	Familie (family)
Gedicht (poem)	Wache (guard)	Aufsicht (supervision)

Running head: Sleep effect in the long range

The long range effect of sleep on episodic memory

Olaf Lahl and Reinhard Pietrowsky

Institute of Experimental Psychology, University of Duesseldorf, Germany

Corresponding Author:

Olaf Lahl
University of Düsseldorf
Institut of Experimental Psychology
Universitätsstr. 1
40225 Düsseldorf
Germany
Phone: ++49-211-811-2146
Fax: ++49-211-8114261
E-mail: olaf.lahl@uni-duesseldorf.de

Abstract

In past studies on the relationship between sleep and episodic memory, enhanced recall performance has usually been observed with short retention intervals of eight hours of sleep or waking. Experiments addressing the question as to whether this effect is also detectable for longer retention intervals of 16 hours up to six days have been scarce and particularly inconclusive. The present study therefore examined free and cued recall of a categorized word list in a 2 × 2 design, completely between subjects, with a period of either sleep or wakefulness following initial learning and the recall test administered either seven or 72 hours later. Results indicated superior memory performance after the long, but not after the short retention interval. The outcome is discussed with respect to the generalizability and robustness of the beneficial effect of sleep on memory.

Key words: sleep, episodic memory, retention interval, long-term effect

Introduction

Today's ongoing debate about a potential involvement of sleep in enhancing memory performance (Stickgold and Walker, 2005; Vertes and Siegel 2005) was originally inspired by early investigations reporting superior episodic recall after interpolated periods of two to eight hours of sleep as opposed to waking (Jenkins and Dallenbach, 1924; Van Ormer, 1932). Subsequent investigations partially confirmed these results for retention intervals of similar length (Benson and Feinberg, 1975, 1977; Ekstrand, 1967; Idzikowski, 1984; Lovatt and Warr, 1968) although other factors such as circadian phase (Hockey *et al.*, 1972; Nesca and Koulack, 1994), different sleep stages (Barrett and Ekstrand, 1972; Plihal and Born, 1997), and sleep cycle integrity (Ficca *et al.*, 2000) soon turned out to play a critical role as well. Due to their limited duration the retention intervals that were used in these studies were either filled with pure sleep or pure waking.

From any presumption of a functional involvement of sleep in memory processing it follows that at least part of the direct sleep memory effect should remain stable over intermediate time periods. However, empirical evidence becomes scarce and less consistent when we turn to the impact that sleep has on episodic memory in the long run. Table 1 summarizes the results of six studies which systematically tested prolonged retention intervals of up to six days following eight hours of post-learning sleep or waking. As may be observed, the overall picture is rather confusing even though nonsense syllables served as learning material in all of the studies but one (Benson and Feinberg, 1977) which used paired associate lists instead.

Insert Table 1 about here

Results for time periods of more than 48 hours are particularly rare as they come from only three studies, two of which allow for only limited conclusions. Graves (1936) found a sleep benefit effect for intervals of 72 to 144 hours, but these data have to be viewed with caution since the investigation was designed as a single case study with the experimenter acting as her own subject. Following the report of Richardson and Gough (1963), Gibb (1937) failed to replicate Graves' results in performing a repeated measures experiment with six subjects. However, not much can be said about the methodological adequacy of his study since the original manuscript has never been published. Almost three decades later in a methodologically sound approach Richardson and Gough (1963) replicated Graves' finding for the 144 hour interval.

Despite such an intricate state of affairs, to our knowledge no other attempt has been made to assess the long term effect of sleep on retention ever since. The present study was therefore designed to compare verbal recall after either post-learning sleep or wakefulness in the short range of 7 h, which served as a control condition, and in the long range of 72 h. For both time intervals we predicted superior retention when acquisition was followed by sleep.

Method

Subjects

Sixty university students (46 female, 14 male) with a mean age of 23.3 ± 4.7 years participated in the experiment in exchange for financial compensation. All subjects were healthy nonsmokers with a regular sleep schedule. They were obliged to refrain from alcoholic beverages, caffeine, and daytime napping beginning 12 hours before and continuing throughout the entire experiment. All participants gave written consent to take part in the study after the experimental protocol had been fully explained.

Memory testing

The learning material consisted of a list of 56 nouns with four words each belonging to one of 14 categories. All items were taken from the category norms provided by Battig and Montague (1969) and were translated into German. To prevent ceiling (guessing) and floor effects during later category-cued recall, only words with category-association ranks between 4-10 were selected for creating the lists.

During learning, the 56 words were presented individually on a computer screen with a rate of one word per 1500 ms and a delay of one second between each word. The complete list was presented twice in succession, with different fixed random orders of words for both sequences. Words were not blocked according to category, nor were the category names presented. The subjects were instructed to pay attention to the presented words and to repeat each word overtly or silently. To prevent subsequent active rehearsal, the learning session was immediately followed by a brief cover task. Subjects had to rate 16 non-aversive pictures taken from the International Affective Picture System (Lang *et al*., 1999) on their perceived level of pleasantness and arousal. Together, the learning and the dummy session took about 10 min.

Delayed memory performance at the end of the different retention intervals was assessed by two separate software-driven recall sessions. During the first test, subjects were given 10 minutes to freely recall as many of the previously presented words as possible. They were instructed to enter the remembered words into an empty text field providing unlimited line space. The second test was a 10 min cued recall test. Fourteen text fields with the category labels serving as cues were presented simultaneously on the screen. Subjects were instructed to assign each remembered word to the correct category. Each text field was limited to four lines and no white space character was accepted as input, so that a maximum of four words could be entered. This way, subjects were not able to produce hits by simply listing all members of a given category that came to their mind.

Design and procedure

The study followed a 2 × 2 factorial design. The first factor was defined by either night sleep or daytime wakefulness subsequent to the initial learning session. The second factor was the overall retention interval of either 7 h or 72 h. Both factors were varied between subjects.

Subjects in the sleep conditions reported to the laboratory at 23.30 h to prepare themselves for retiring. Thereafter, they underwent the learning session and were put to bed with lights turned off immediately. After seven hours, they were awakened by the experimenter.

Subjects with a 7 h retention interval performed the recall session immediately thereafter. Those with a 72 h retention interval were dismissed from the laboratory and returned for recall three days later at 23.30 h, i.e. 72 h after initial learning. In both wake conditions, subjects had their original learning at 08.30 h and were subsequently dismissed from the laboratory to follow their usual daytime activities. They returned for recall the same day at 15.45 h (7 h condition) or three days later at 08.30 h (72 h condition).

Results

By visual inspection, subjects in both sleep conditions reproduced more words on the free recall task (Figure 1A) as well as on the categorized recall task (Figure 1B) than their wake counterparts. The average (mean ± standard error of mean) free recall rate after seven hours was 20.53 ± 2.46 in the sleep condition and 18.33 ± 2.04 in the wake condition. As expected, retention declined markedly over time, yielding free recall rates of 13.93 ± 1.78 and 9.67 ± 1.74 in the 72 h sleep and wake condition respectively. Cueing by category labels improved recall in all groups to an overall extent of 5.10 items on average. Cued recall was 25.86 ± 2.94 and 23.27 ± 2.11 after seven hours of sleep and waking and 19.20 ± 1.93 and 15.40 ± 2.18 for the respective 72 h retention conditions.

Insert Figure 1 about here

The product-moment correlation between free recall and cued recall was $r = 0.95$, so that both scores virtually measured the same property. We therefore restricted the inferential statistics to an univariate analysis of the free recall measure. Pairwise planned comparisons (sleep vs. wake after 7h, sleep vs. wake after 72 h) using one-tailed *t* tests and the pooled mean error variance ($MSE = 61.52$) revealed an insignificant effect of sleep for the 7 h retention interval [$t(56) = 0.77$; $P = .223$] as well as for the 72 h retention interval [$t(56) = 1.49$; $P = .071$]. Note, that the latter comparison reaches statistical significance when the overall error variance is replaced by the individual error variances of the two involved groups (sleep / 72 h: $s^2 = 47.75$; wake / 72h: $s^2 = 45.56$) in the calculation of the *t* statistic [$t(28) = 1.71$; $P = .049$].

Discussion

With the present study we aimed to clarify whether the beneficial effect of sleep on retention is stable over a prolonged time period of 72 hours. The seven hour retention interval was included as a control condition, but contrary to predictions, no short term sleep advantage could be demonstrated. Regarding memory performance after 72 hours, the statistical inference is not definite, but does suggest a beneficial effect from sleep over a longer time frame. This pattern of results – sleep becoming effective only after a certain time span has elapsed – is similar to the findings obtained by Graves (1936) and Richardson and Gough (1963), but stands in contrast to the outcome of Gibb’s (1937) experiment.

Turning back to Table 1, the situation somewhat resembles the state of affairs concerning rapid eye movement (REM) sleep windows. These are supposed to represent critical episodes of increased REM sleep following exposure to learning demands which, when disrupted, impair memory consolidation of the learned task. While the basic idea is straightforward, the moving character of REM windows – their reported time of occurrence ranging between one and 56 hours after learning (Smith, 1995) – is an issue of major concern (Vertes and Eastman, 2000).

Although from a theoretical viewpoint the assumed time dependency of a sleep-related memory facilitation is difficult to explain, consolidation theory might offer a preliminary account. Consolidation theory proposes the gradual shift of newly acquired memories from an initially volatile state to a stabilized memory trace which is less fragile and less sensitive to disruption. Since the precise time course of the underlying process appears to be a free parameter of the theory (Meeter and Murre, 2004), it may be speculated that post-learning sleep offers optimal initial conditions for triggering consolidation but that the entire process in fact continues far beyond that period. Thus, the effectiveness of the initial sleep period would not become obvious until complete termination of the process. Of course, this approach remains somewhat fuzzy as long as the precise duration of the presumed sequence of events is not specified.

As was outlined in the introduction, sleep has been shown many times to enhance episodic recall in a short time frame. However, the relationship between sleep and episodic memory seems to be neither simple nor robust (Rauchs *et al.*, 2005; Walker and Stickgold, 2004). Other factors appear to contribute equally or even more in determining what is retained and what is not, and altering one of them is likely to change the overall picture. Accordingly, several investigators found no effect of sleep on episodic memory when they systematically varied factors such as circadian phase (Hockey *et al.*, 1972; Nesca and Koulack, 1994), amount of rapid eye movement sleep (Yaroush *et al*., 1971), emotional salience (Wagner, Gais and Born, 2001) and significance of the learning material (Newman, 1939), and level of cerebral acetylcholine (Gais and Born, 2004). The time span between acquisition and retrieval might be yet another factor contributing to the complicated relationships.

On the other hand, we know of not a single study reporting superior retention after a period of waking as opposed to sleep. To draw a tentative conclusion, we would like to suggest that much of the empirical evidence actually speaks in favor of a sleep-related memory protection to some degree. We feel, however, that the effect is rather incidental and has clear limitations in terms of magnitude and robustness and therefore probably does not address the – yet unknown – primary function(s) of sleep.

References

Barrett, T. R. and Ekstrand, B. R. Effect of sleep on memory III. Controlling for time-of-day effects. *J. Exp. Psychol.*, 1972, 96: 321-327.

Battig, W. F. and Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *J. Exp. Psychol. Monograph* 1969, 80.

Benson, K. and Feinberg, I. Sleep and memory: retention 8 and 24 hours after initial learning. *Psychophysiology*, 1975, 12: 192-195.

Benson, K. and Feinberg, I. The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology*, 1977, 14: 375-384.

Ekstrand, B. R. Effect of sleep on memory. *J. Exp. Psychol.*, 1967, 75: 64-72.

Ficca, G., Lombardo, P., Rossi, L. and Salzarulo, P. Morning recall of verbal material depends on prior sleep organization. *Behav. Brain Res.*, 2000, 112: 159-163.

Gais, S. and Born, J. Low acetylcholine during slow wave sleep is critical for declarative memory consolidation. *P. Natl. Acad. Sci. USA*, 2004, 101: 2140-2144.

Graves, E. A. The effect of sleep upon retention. *J. Exp. Psychol.*, 1936, 19: 316-322.

Hockey, G. R. J., Davies, S. and Gray, M. M. Forgetting as a function of sleep at different times of day. *Q. J. Exp. Psychol.*, 1972, 24: 386-393.

Idzikowski, C. Sleep and memory. *Br. J. Psychol.*, 1984, 75: 439-449.

Jenkins, G. and Dallenbach, K. M. Obliviscence during sleep and waking. *Am. J. Psychol.*, 1924, 35: 605-612.

Lang, P. J., Bradley, M. M. and Cuthbert, B. N. *The international affective picture system (IAPS): digitized photographs*. The Center of Research in Psychophysiology, University of Florida, 1999.

Lovatt, D. J. and Warr, P. B. Recall after sleep. *Am. J. Psychol*., 1968, 81: 253-257.

Meeter, M. and Murre, J. M. J. Consolidation of long-term memory: Evidence and alternatives. *Psychol. Bull*., 2004, 130: 843-857.

Nesca, M. and Koulack, D. Recognition memory, sleep and circadian rhythms. *Can. J. Exp. Psychol*., 1994, 48: 359-379.

Newman, E. B. Forgetting of meaningful material during sleep and waking. *Am. J. Psychol.*, 1939, 52: 65-71.

Plihal, W. and Born, J. Effects of early and late nocturnal sleep on declarative and procedural memory. *J. Cogn. Neurosci*., 1997, 9: 534-547.

Rauchs, G., Desranges B., Foret, J. and Eustache, F. The relationship between memory systems and sleep stages. *J. Sleep Res.*, 2005, 14: 123-140.

Richardson, A. and Gough, J. E. The long range effect of sleep on retention. *Aust. J. Psychol.*, 1963, 15: 37-41.

Smith, C. Sleep states and memory processes. *Behav. Brain Res.*, 1995, 69: 137-145.

Stickgold, R. and Walker, M. P. Sleep and memory: The ongoing debate. *Sleep*, 2005, 28: 1225-1227.

Van Ormer, E. B. Retention after intervals of sleep and of waking. *Arch. Psychol*., 1932, 137: 1-49.

Vertes, R. P. and Eastman, K. E. The case against memory consolidation in REM sleep. *Behav. Brain Sci.*, 2000, 23: 867-876.

Vertes, R. P. and Siegel, J. M. Time for the sleep community to take a critical look at the purported role of sleep in memory processing. *Sleep*, 2005, 28: 1228-1229.

Walker, M. P. and Stickgold, R. Sleep-dependent learning and memory consolidation. *Neuron*, 2004, 44: 121-133.

Wagner, U., Gais, S. and Born, J. Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learn. Mem.*, 2001, 8: 112-119.

Yaroush, R., Sullivan, M. J. and Ekstrand, B. R. Effects of sleep on memory II: differential effect of the first and second half of the night. *J. Exp. Psychol.*, 1971, 88: 361-366.

Tables

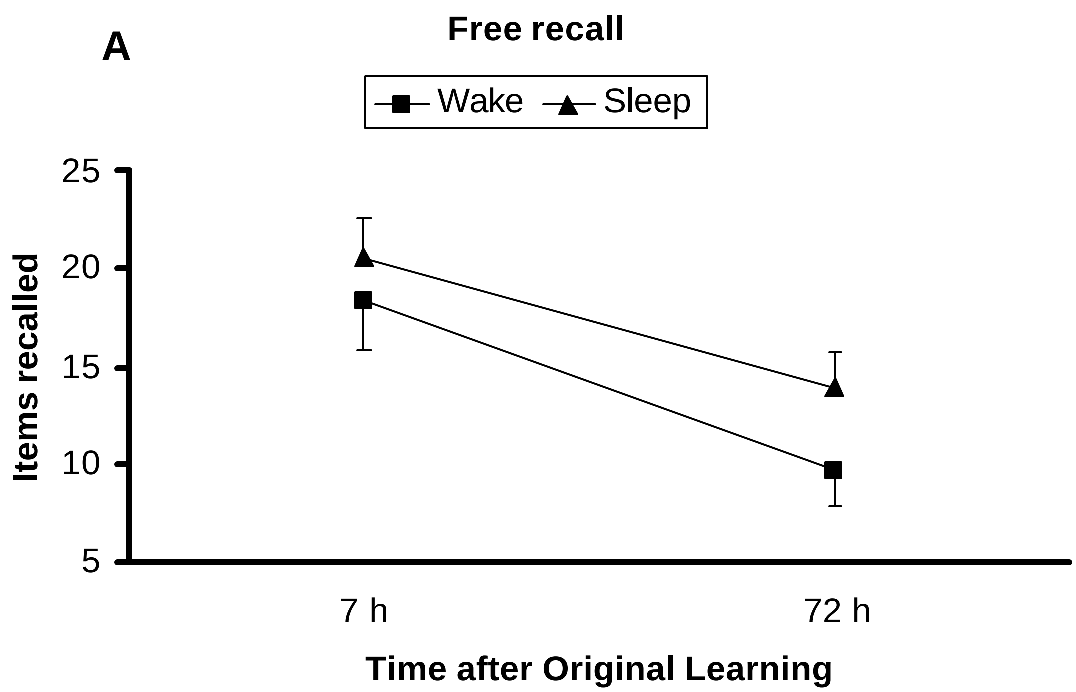
Table 1: Results of studies comparing recall as a function of post learning sleep or waking after different retention intervals

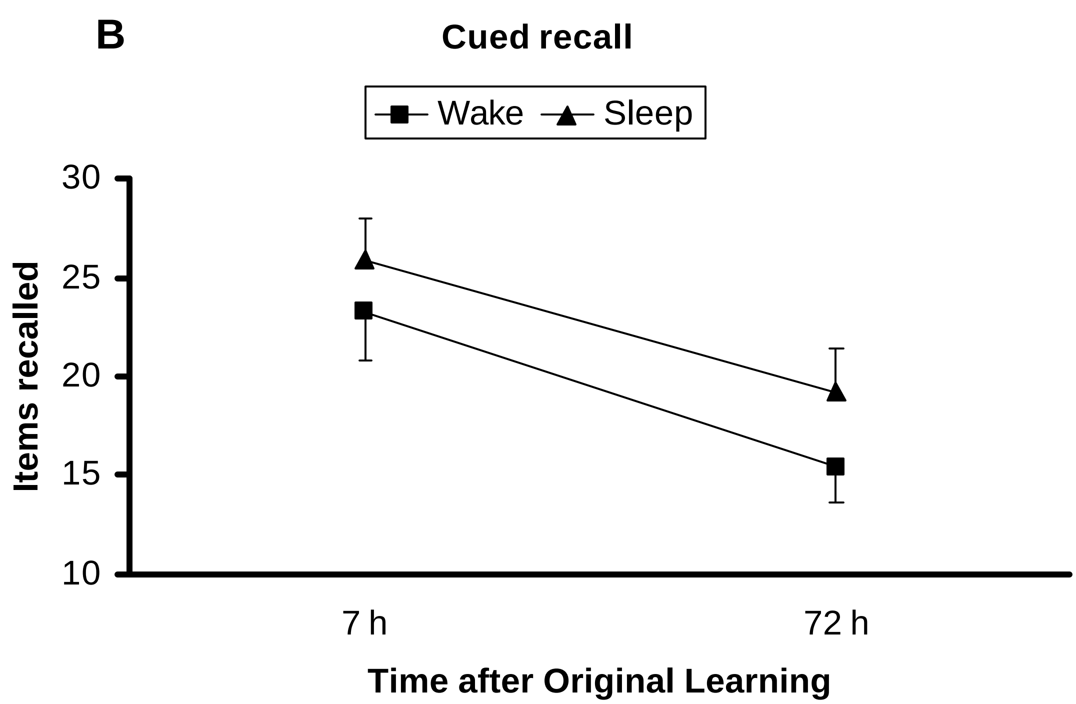
	Retention interval [h]					
Study	16	24	48	72	96	144
Benson and Feinberg (1975)		ø				
Benson and Feinberg (1977)	ø	+				
Gibb (1937)		ø	+	ø	ø	
Graves (1936)		ø	ø	+	+	+
Idzikowski (1984)	+	+				
Richardson and Gough (1963)		ø	ø			+

Note. +: superior recall after sleep, ø: no difference between recall after sleep and waking.

Figure Captions

Figure 1. Number of words (*M* ± *SEM*) retained after retention intervals of 7 h and 72 h when original learning was followed by either sleep or wakefulness. **A**. Free recall. **B** Cued recall.


A
Free recall
Wake
Sleep
Items recalled
25
20
15
10
5
7 h
72 h
Time after Original Learning


B
Cued recall
Wake
Sleep
Items recalled
30
25
20
15
10
7 h
72 h
Time after Original Learning

Effect of sleep on memory for task-irrelevant stimuli

Olaf Lahl and Reinhard Pietrowsky

Institute of Experimental Psychology, University of Duesseldorf, Germany

Corresponding Author:

Olaf Lahl
Universität Düsseldorf
Klinische Psychologie
Universitätsstr. 1 (Geb. 23.03)
40225 Düsseldorf
Germany
Phone: ++49-211-811-2146
Fax: ++49-211-8114261
E-mail: olaf.lahl@uni-duesseldorf.de

Effect of sleep on memory for task-irrelevant stimuli

Abstract

Cumulative evidence suggests that sleep has a facilitative effect on the retention of rote learning material. However, little is known about its impact on task-irrelevant material that was acquired incidentally. This study examined a verbal and a visual task of incidental memory after 7 h of nocturnal sleep vs. nocturnal wakefulness (sleep deprivation) and diurnal wakefulness. For the verbal task subjects had to form meaningful sentences that related word pairs under the pretense of conducting a creativity test. For the visual task subjects were instructed to memorize a series of pictures, each of which could appear in one of the four corners of the screen. Unexpected to subjects, the delayed tests prompted for cued recollections of word pairs and picture locations. Recall for word pairs did not differ significantly between the three groups. Recall for picture locations was significantly better after sleep than after nocturnal wakefulness, but no significant difference was obtained when sleep was compared to daytime wakefulness. The results indicate that the beneficial effect sleep has on explicitly learned material does not extend to the domain of incidental memory.

Key words: incidental learning, unlearning, sleep, retention, task-relevance

Introduction

For 80 years now, since Jenkins and Dallenbach published their pioneering paper (Jenkins & Dallenbach, 1924) about "obliviscence during sleep and waking", the putative role of sleep in enhancing learning and memory has been heavily debated in the scientific community (Stickgold & Walker, 2005; Vertes, 2004; Vertes & Siegel, 2005; Walker & Stickgold, 2004). Superior retention after periods of sleep as opposed to wakefulness was subsequently confirmed for various verbal learning materials, including nonsense syllables (Benson & Feinberg, 1975; Idzikowski, 1984; Van Ormer, 1932), lists of single words (Hockey, Davies, & Gray, 1972; Nesca & Koulack, 1994), paired associate lists (Benson & Feinberg, 1977; Ekstrand, 1967; Ekstrand, Barrett, West & Maier, 1977; Lovatt & Warr, 1968; Spight, 1928), and short stories (Newman, 1939). A common feature of this early branch of research was the devotement to explicit rote learning. That is, at the time of encoding subjects were fully aware of the fact that their memory was of primary relevance in the ongoing experiment, be it to reach a certain learning criterion or in anticipation of a forthcoming test of retention. Subjects therefore deliberately attempted to memorize as much of the task-relevant material as possible.

In contrast, "learning is implicit when it is unaffected by intention" (Stadler & Frensch, 1994, p. 423). In implicit tasks, the intention to acquire and memorize material is typically eliminated by a cover task such as rating words on their pleasantness or counting their letters, when in fact these words are to be recalled later. Hence, the learning process is said to have happened en passant, or *incidentally*.

To our knowledge, few studies on sleep and memory have deployed implicit learning material. Plihal and Born (1999) used a wordstem priming paradigm to assess differential effects of early night sleep, which is dominated by the occurrence of slow wave sleep (SWS), versus late night sleep, which is prevailed by the occurrence of stage two and rapid eye movement (REM) sleep. Four groups of subjects implicitly learned a list of nouns by rating their melody and were prompted for wordstem completion after 3 h of sleep or wakefulness throughout the first (23.00 h – 02.00 h) or second (03.00 h – 06.00 h) half of the night. Implicit recall was better in the late sleep condition than in the late wake condition, but no significant difference was found for the respective early night periods. Wagner, Hallschmid, Verleger, and Born (2002) applied the same experimental design to faces that were learned implicitly by instructing subjects to indicate the faces' sex. A memory advantage in terms of decreased reaction times was again found for the late sleep condition when subjects had to judge the familiarity of the faces. Although both studies tap into the mechanisms of implicit memory, they share the confinement to fractional sleep. Moreover, as far as the second experiment is concerned, face memory is supposed to represent a distinct memory system which operates differently from and independently of others (Farah, 1996).

In the present study we were interested in the effect that uninterrupted sleep has on the retention of stimuli that were task-irrelevant to the subject at the time of acquisition and thus learned incidentally. Our interest in task-irrelevant material was partially inspired by the proposal of processes of *unlearning* or reverse learning during sleep (Crick & Mitchison, 1983, 1995; Evans & Newman, 1964; Newman & Evans, 1965). In an attempt to provide a theory for the function of REM sleep, Crick and Mitchison hypothesized that this sleep stage serves to clear cortical neural nets from "unwanted or 'parasitic' modes of behavior" (Crick & Mitchison, 1983,

p. 111) thereby preventing overload and making “the storage in an associative net more efficient” (Crick & Mitchison, 1995, p. 149).

While this theory “is essentially a micro theory” (Crick & Mitchison, 1995, p. 155) formulated in terms of neural network theory, we were mainly interested in the basic concept of sleep in order to forget or reject “redundant or inapposite memories and responses” (Evans & Newman, 1964, p. 577). Based on the assumption of limited retrieval capacities in human memory (Anderson & Schooler, 1991) and further presuming that sleep plays an essential role in optimizing overall memory performance, a weakening of memory for task-irrelevant material that is unlikely to be needed in a future task would be expected. We therefore compared the effects of a full night of sleep with the effect of nocturnal wakefulness (sleep deprivation) and diurnal wakefulness on the retention of incidentally acquired semantic and visual learning material.

Methods

Participants

Fifty university students (38 female, 12 male) aged between 19-29 years (M = 23.76, SD = 2.66) participated in the experiment in exchange for financial compensation and were allocated randomly to either the Sleep/Night (n = 15), Wake/Night (n = 15), or Wake/Day (n = 20) condition. All subjects were nonsmokers, had a regular nightly sleep schedule, and were free of any psychiatric illness and psychoactive medication. They were obliged to refrain from alcoholic beverages, caffeine, and daytime napping 12 h before embarking on the experiment. All participants gave signed consent to take part in the study after the study protocol had been fully explained.

Dependent variables

The test paradigm for incidental memory of verbal material was similar to the task employed in [15]. A computer program presented 16 pairs of semantically unrelated German nouns (for example *industry–pleasure*), one at a time. The subjects' task was to type a meaningful sentence for each word pair (e.g., *Working in the industry is of much pleasure*) into a text field within 45 sec, ostensibly in order to "probe their capability of creative thinking." Thus, subjects were led to believe that the task was completed once all sentences were created. At the end of the retention interval, however, they were given an unexpected cued recall test for the response words. Only the stimulus words were then presented sequentially by the software and subjects were required to complete each pair by typing the matching response word into a text field. Subjects were not given a time limit for this task. Since a single recall trial is likely to underestimate memory performance (Brown, 1923; Lazar, 1969; Lazar & Van Laer, 1966; Payne, 1986; Richardson & Gropper, 1964), subjects were given three successive recall trials, each with a different order of stimulus words.

In order to assess incidental memory for visual stimuli, we adopted a standard central-incidental paradigm that has been successfully utilized for this purpose in previous studies (Andersson & Hockey, 1977; Davies & Jones, 1975; Hockey & Hamilton, 1970; Jubis, 1990). Twenty-four neutral photographs of everyday objects were taken from the International Affective Picture System (Lang, Bradley, & Cuthbert, 1999). The computer software presented each picture individually for 2000 ms with an inter-stimulus-interval of 1000 ms. Each picture could appear in one of the four corners of the screen, with the restriction that all locations occurred with equal frequency and no successive locations were the same. Subjects were requested to memorize the pictures with respect to a supposed forthcoming test of retention for those pictures (task-relevant component). No mention was made regarding the relevance of the screen locations (task-irrelevant component), which in fact had to be recalled at the end of the 7 h retention interval.

In order to prevent subjects from suspecting the true purpose of this task, they had to accomplish two dummy tasks up front. First, a set of eight two-digit numbers was presented in exactly the same manner as the photographs with each number appearing in one corner of the screen. Subjects were instructed to memorize those numbers for an immediate test of retention, *which did in fact occur*. Subsequently, the same procedure of learning with immediate recall was repeated for a set of 16 common forenames. Thus, when subjects finally underwent the actual incidental learning session for the set of photographs, they would have already experienced two test sessions which met their expectations. Furthermore, they would have observed stimuli appearing at different screen corners without any apparent significance to the ongoing task two times. Consequently, when subjects were studying the pictures after all, they had no reason to expect anything other than a delayed recall test for this very set of pictures.

At the end of the retention interval, the pictures were individually displayed in the middle of the screen (with the presentation order being different from the one used during learning). The task was to indicate the original screen location of each picture by clicking one of four push buttons that were positioned below the picture. These were arranged in a 2 × 2 grid so that each buttons' position within the grid matched the screen location it represented.

Design and procedure

For both night groups the retention interval was the nocturnal period between 00.00 h and 07.00 h, for the daytime wake group it was the period between 09.00 h and 16.00 h. The duration of seven hours was chosen with regard to recent research demonstrating that this is the average habitual sleep period of healthy young adults (Breslau, Roth, Rosenthal, & Andreski, 1997; Hicks, Lucero-Gorman, Bautista, & Hicks, 1999; Hirshkowitz, Moore, Hamilton, Rando, & Karacan, 1992). The learning session was scheduled for 23.30 h (night groups) and 08.30 h (day group) respectively and always lasted about 25 min. The recall session was scheduled for 07.00 h (night groups) and 16.00 h (day group) respectively and took between 10 and 20 min; the precise duration depending on individual response speed. At the end of the experiment all participants had to indicate whether they had anticipated the recall tests for word pairs and picture locations.

In the sleep condition, sleep was polysomnographically monitored according to standard criteria (Rechtschaffen & Kales, 1968) during the entire night. Sleep Subjects were acclimated to the placement of electrodes and the sleep laboratory by spending one adaptation night. To ensure comparability with the wake conditions, the adaptation night was always scheduled two nights before the test night so that participants of all groups would sleep at home the night before the experiment. On the experimental night, subjects of the sleep group reported to the laboratory at 22.30 h to prepare themselves for retiring and to have their electrodes applied. When they had finished their learning session, they were put to bed immediately and the lights were turned off. Subjects were awakened seven hours later. They were allowed to remove the electrodes and to dress themselves before starting the recall session.

Subjects in the nightly wake condition remained awake between learning and recall and were under permanent control of always two experimenters. Subjects in the daytime wake condition were dismissed from the laboratory after they had finished the learning session. They were obliged to refrain from daytime napping and again reported to the laboratory for delayed recall. To prevent any exchange of hypotheses about the purpose of the experiment, subjects of all groups participated individually.

Data analysis

Sleep recordings were analyzed off-line according to standard criteria (Rechtschaffen & Kales, 1968). Relevant sleep parameters were sleep onset latency, amounts of sleep stages 1, 2, SWS, and REM as well as total sleep time in relation to time spent in bed (sleep efficiency).

Word pair recall was assessed cumulatively over all three successive recall trials. A correct response was thus scored the first time it was produced while ignoring subsequent repetitions as well as subsequent failures to reproduce the same item again. Memory performance for visual stimuli was defined by the number of picture locations correctly recalled.

Both recall measures were tested between conditions Wake/Night vs. Sleep/Night and Wake/Day vs. Sleep/Night using Dunnett's two-tailed *t* test for comparing all treatments with a control (Dunnett, 1955). Dunnett's test ensures that the family-wise Type-I error rate for multiple comparisons with a single control (i.e. the sleep condition) does not exceed the predefined level of $\alpha = 0.05$. The according test statistics will be referred to as t_d.

Results

Sleep parameters in the experimental night (Table 1) indicated good sleep quality in terms of sleep physiology and thereby successful realization of the sleep condition.

Insert Table 1 about here

Two participants in the Sleep/Night condition and three participants in the Wake/Day condition had suspected the recall test for word pairs. Their data was omitted in the analysis of verbal recall. None of the subjects had anticipated the recall test for picture locations, so the entire data sample was included in this analysis.

Figure 1 shows the cumulative recall curves over the three successive recall trials (designated as T1, T2 cum., T3 cum.) for word pairs. As can be seen, the number of completed word pairs increased slightly over trials in all groups. However, performance (M ± SD) in the sleep condition was higher than in both wake conditions at all stages of testing. Specifically, cumulative recall was higher after sleep (13.23 ± 1.69) than after nocturnal wakefulness (10.60 ± 3.94) and diurnal wakefulness (11.00 ± 3.14). The comparison Wake/Night vs. Sleep just failed to reach significance [$t_d(47) = 2.22$; $p = .057$]. Comparing Wake/Day with Sleep also yielded a non-significant result [$t_d(47) = 1.94$; $p = .10$].

Insert Figure 1 about here

Picture location recall (Figure 2) was well above the binomial 5% chance level of ten pictures in all groups (Sleep: 17.07 ± 4.01; Wake/Night: 13.20 ± 3.36; Wake/Day: 15.30 ± 3.97). As with verbal recall, subjects in the sleep group outperformed both control groups. This superiority was significant for the comparison Wake/Night vs. Sleep [$t_d(47) = 2.78$; $p = .015$], but not for the comparison Wake/Day vs. Sleep [$t_d(47) = 1.36$; $p = .301$].

Insert Figure 2 about here

Discussion

This study assessed memory for task-irrelevant verbal and visual stimuli after conditions of sleep during the night and wakefulness during the night and during the day. Contrary to the assumption of better removal (lower recall) of those stimuli after sleep due to processes of memory clearance, recall rates were even higher after sleep than after waking. For the visual task this superiority of the sleep condition was significant when compared to sleep deprivation. For the verbal task the same comparison just failed to reach significance. When sleep was compared to daytime wakefulness, none of the two recall measures provided a statistically reliable result. We therefore cannot exclude the possibility that the lower recall rates in the sleep deprivation group were in part due to unspecific effects of fatigue.

Much experimental effort was spent on the operationalization of task-irrelevance. The very low rate of subjects who had anticipated the delayed recall tests confirm that the learned material was truly out of subjects' focus of attention during the critical periods of acquisition and retention. Accordingly, the assumption of "hidden" explicit learning processes cannot account for the negative results.

In summary, the present study clearly does not favor the notion of sleep as a selective memory enhancer that filters relevant from irrelevant data. Although such a filter function seems to have much intuitive and theoretical appeal, the current data suggest no effect or even a facilitation of memory for extraneous pieces of information.

One apparent approach to explain these negative results is to assume a delay in the distinction between informational units of high and low relevance. In that case all elements would initially be designated as potentially relevant. Only in the *long range* of time would the supposed process of making "most available those items that were most likely to be needed" (Anderson & Schooler, 1991, p. 396) while dampening the others occur. Previous evidence suggests that the longer course of time indeed may be of crucial importance. When conditions of post-learning sleep was compared with post-learning wakefulness after retention intervals of 24, 48, and 144 hours, a significant memory advantage for the sleep condition was found only after the time period of 144 hours (Richardson & Gough, 1963). A reverse effect may exist for irrelevant pieces of information. Future research is challenged to test this reformulation of the unlearning account.

References

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, **2**, 396-408.

Andersson, K., & Hockey, G. R. J. (1977). Effects of cigarette smoking on incidental memory. *Psychopharmacology*, **52**, 223-226.

Benson, K., & Feinberg, I. (1975). Sleep and memory: retention 8 and 24 hours after initial learning. *Psychophysiology*, **12**, 192-195.

Benson, K., & Feinberg, I. (1977). The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology*, **14**, 375-384.

Breslau, N., Roth, T., Rosenthal, M. D., & Andreski, M. A. (1997). Daytime sleepiness: An epidemiological study of young adults. *American Journal of Public Health*, **87**, 1649-1653.

Brown, W. (1923). To what extent is memory measured by a single recall? *Journal of Experimental Psychology*, **6**, 377-382.

Crick, F., & Mitchison, G. (1983). The function of dream sleep. *Nature*, **304**, 111-114.

Crick, F., & Mitchison, G. (1995). REM sleep and neural nets. *Behavioral Brain Research*, **69**, 147-155.

Davies, D. R., & Jones, D. M. (1975). The effects of noise and incentives upon attention in short-term memory. *British Journal of Psychology*, **66**, 61-68.

Dunnett, C.W. (1955). A multiple comparison procedure for comparing several treatments with a control, *Journal of the American Statistical Association*, **50**, 1096-1121.

Ekstrand, B. R. (1967). Effect of sleep on memory. *Journal of Experimental Psychology*, **75**, 64-72.

Ekstrand, B. R., Barrett, T. R., West, J. N., & Maier, W. G. (1977). The effect of sleep on human long-term memory. In R. R. Drucker-Colin, J. L. McGaugh (Eds.), *Neurobiology of sleep and memory* (pp. 419-438). New York: Academic Press.

Evans, C. R., & Newman E. A. (1964). Dreaming: an analogy from computers. *New Scientist*, **419**, 577-579.

Farah, M.J. (1996). Is face recognition 'special'? Evidence from neuropsychology. *Behavioral Brain Research*, **76**, 181-189.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Human Learning & Memory*, **11**, 501-518.

Hicks, R. A., Lucero-Gorman, K., Bautista, J., & Hicks G. J. (1999). Ethnicity, sleep duration, and sleep satisfaction. *Perceptual & Motor Skills*, **88**, 234-235.

Hirshkowitz, M., Moore, C. A., Hamilton, C. R, Rando, K. C., & Karacan, I. (1992). Polysomnography of adults and elderly: Sleep architecture, respiration, and leg movement. *Journal of Clinical Neurophysiology*, **9**, 56-62.

Hockey, G. R. J., Davies, S., & Gray, M. M. (1972). Forgetting as a function of sleep at different times of day. *Quarterly Journal of Experimental Psychology*, **24**, 386-393.

Hockey, G. R. J., & Hamilton, P. (1970). Arousal and information selection in short-term memory, *Nature*, **226**, 866-867.

Idzikowski, C. (1984). Sleep and memory. *British Journal of Psychology*, **75**, 439-449.

Jenkins, J. G., & Dallenbach, K. M. (1924). Obliviscence during sleep and waking. *American Journal of Psychology*, **35**, 605-612.

Jubis, R. M. T. (1990). Effects of alcohol and white noise on recall of relevant and irrelevant task components. *Perceptual & Motor Skills*, **71**, 691-702.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *The international affective picture system (IAPS): digitized photographs.* The Center of Research in Psychophysiology, University of Florida.

Lazar, G. (1969). Retention as a function of successive recall trials following original leaming. *Psychological Reports*, **25**, 567-574.

Lazar, G., & Van Laer, J. (1966). Successive recall as a warm-up for paired adjectives. *Psychonomic Sciences*, **5**, 137-138.

Lovatt, D. J., & Warr, P. B. (1968). Recall after sleep. *American Journal of Psychology*, **81**, 253-257.

Nesca, M., & Koulack D. (1994). Recognition memory, sleep and circadian rhythms. *Canadian Journal of Experimental Psychology*, **48**, 359-379.

Newman, E. B. (1939). Forgetting of meaningful material during sleep and waking. *American Journal of Psychology*, **52**, 65-71.

Newman, E. A., & Evans, C. R. (1965). Human dream processes as analogous to computer programme clearance. *Nature*, **206**, 534.

Payne, D. G. (1986). Hypermnesia for pictures and words: Testing the recall level hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **12**, 16-29.

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

Rechtschaffen, A., & Kales, A. (1968). *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects.* Washington: Public Health Service Publications.

Richardson, A., & Gough, J. E. (1963). The long range effect of sleep on retention. *Australian Journal of Psychology*, **15**, 37-41.

Richardson, J., & Gropper, M. S. (1964). Learning during recall trials. *Psychological Reports*, **15**, 551-560.

Spight, J. B. (1928). Day and night intervals and the distribution of practice. *Journal of Experimental Psychology*, **11**, 397-398.

Stadler, M. A., & Frensch, P. A. (1994). Whither learning, whither memory? *Behavioral Brain Sciences*, **17**, 423-424.

Stickgold, R., & Walker, M. P. (2005). Sleep and memory: The ongoing debate. *Sleep*, **28**, 1225-1227.

Van Ormer, E. B. (1932). Retention after intervals of sleep and waking. *Archives of Psychology*, **137**, 1-49.

Vertes, R. P. (2004). Memory consolidation in sleep: Dream or reality. *Neuron*, **44**, 134-148.

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. *Sleep*, **28**, 1228-1229.

Wagner, U., Hallschmid, M., Verleger, R., & Born, J. (2002). Signs of REM sleep dependent enhancement of implicit face memory: a repetition priming study. *Biological Psychology*, **62**, 197-210.

Walker, M. P., & Stickgold, R. (2004). Sleep-dependent learning and memory consolidation. *Neuron*, **44**, 121-133.

Tables

Table 1: Sleep parameters in the experimental night

Sleep parameter	*Mean*	*SD*
Time in bed [min]	408.63	7.02
S1-Sleep onset latency [min]	14.47	11.85
S2-Sleep onset latency [min]	20.77	13.35
Sleep efficiency [%]	94.50	4.26
% Wake	5.50	4.26
% S1	7.27	3.79
% S2	51.99	4.35
% SWS	19.53	3.87
% REM	14.04	4.39
% MT	1.67	0.84

Note. S1: sleep stage 1; S2: sleep stage 2; SWS: slow wave sleep; REM: rapid eye movement sleep; MT: movement time. Percentages are relative to time in bed.

Figure Captions

Figure 1. Number of word pairs recalled over three successive recall trials after seven hours of nocturnal sleep, nocturnal wakefulness, or diurnal wakefulness. Recall on the second and third trial is scored cumulatively with a correct response being scored the first time it is produced while ignoring subsequent repetitions or failures. Error bars indicate standard error of mean.

Figure 2. Number of picture locations recalled after seven hours of nocturnal sleep or wakefulness. Note that any recall score above nine is beyond the binomial 5% chance level which is indicated by the horizontal line. Error bars indicate standard error of mean.

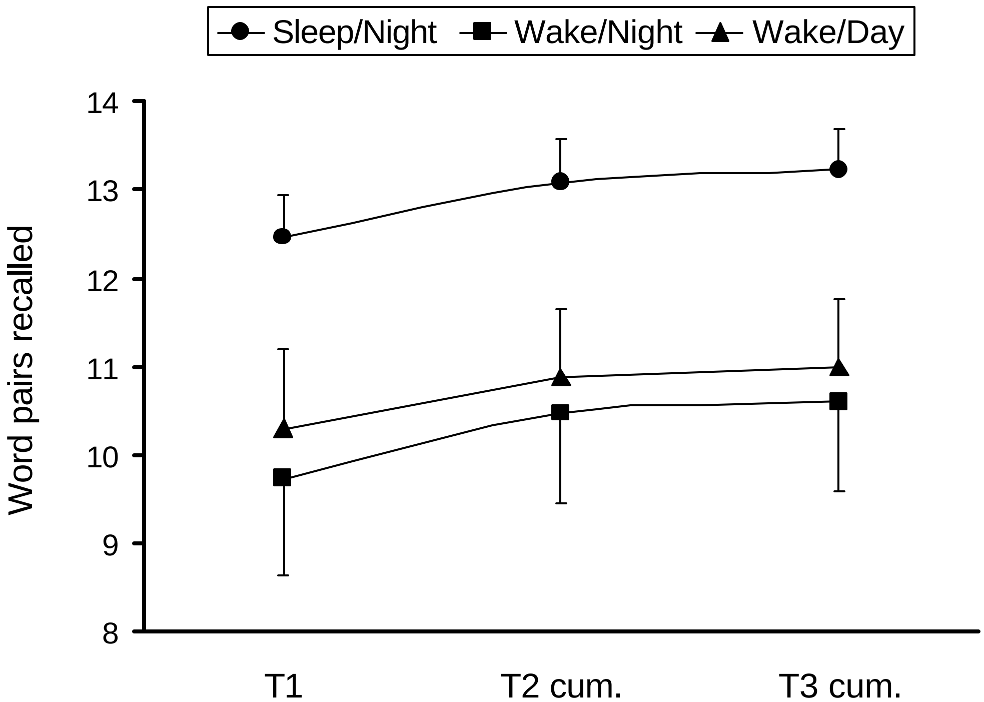

Sleep/Night
Wake/Night
Wake/Day
Word pairs recalled
14
13
12
11
10
9
8
T1
T2 cum.
T3 cum.

Figure 1

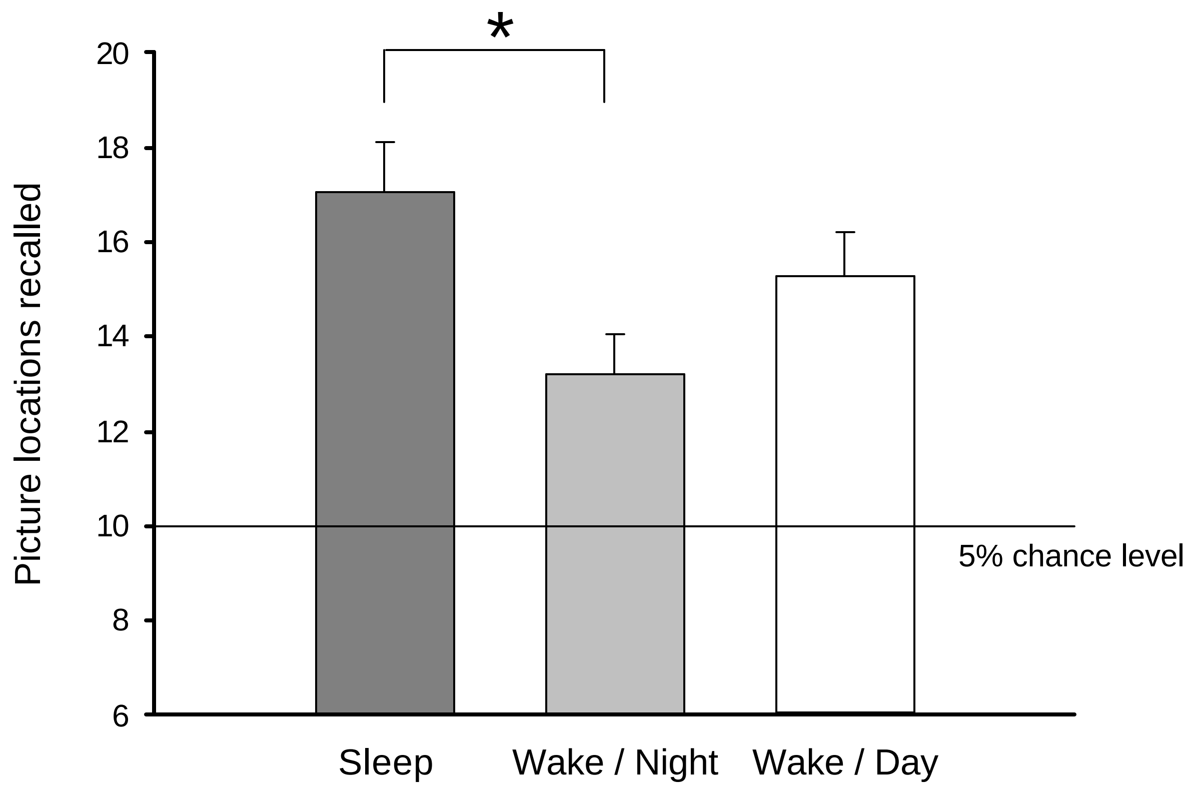

*
20
18
16
14
12
10
8
6
Picture locations recalled
5% chance level
Sleep
Wake / Night
Wake / Day

Figure 2

Sleep enhances memory for things that never happened

Olaf Lahl, Armin Zlomuzica & Reinhard Pietrowsky

Institute of Experimental Psychology, University of Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

In recent years sleep has been shown to facilitate memory consolidation. This effect has been established for a variety of tasks including perceptual[1-3] and motor[4,5] skills, episodic and semantic memory[6-11], and complex cognitive tasks such as gaining insight into hidden rules or schemas[12]. While standard memory tasks typically focus on verbatim learning, requiring subjects to reproduce exactly what was presented during the acquisition phase (e. g. a list of words), psychologists know for a long time that human memory usually does not provide an exact replay of what has happened. Instead, the process of remembering is highly reconstructive in nature and therefore subject to systematic distortions and intrusions[13-17]. Here we extend the knowledge about sleep and memory to the domain of memory for such intrusions or false memories. Subjects' tendency to falsely recognize words not presented during learning was dramatically higher after a night of sleep than after a period of normal daytime wakefulness. We conclude that the beneficial effect of sleep on remembering things is not dependent on whether these things actually occurred, or were created by the subject.

Based on the early work of Deese[18], Roediger and McDermott[19] invented an experimental paradigm – now called the DRM paradigm – that makes it easy to routinely study false memories in the laboratory. In this paradigm, subjects listen to lists of words (e.g. candy, sugar, taste, nice, honey, …) which are all semantically related to a critical theme word (sweet) which is *not* presented. Then, in subsequent memory tests subjects strongly tend to falsely recall or recognize these *critical lures*[19-22]. Furthermore, these false reproductions occur with high subjective confidence and are frequently accompanied by subjects' reports of mentally reliving the experience of having studied the non-presented theme words.

To investigate the effect that sleep has on false memories we had two groups of subjects listen to a set of 18 DRM word lists, each of which contained 15 semantic associates of a critical theme word which was not given. Recognition performance for presented and non-presented words was assessed after a seven hour retention interval that was either filled with nocturnal sleep (sleep group, *n* = 13) or diurnal wakefulness (wake group, *n* = 13). On the delayed recognition task subjects had to classify six items per list as either previously studied or unstudied: (i) three words from each list that had actually been presented during learning, (ii) two semantically unrelated distractor words, and (iii) the critical theme word, with the latter three not having been presented during learning.

Whereas the proportions of correctly recognized list items and falsely recognized unrelated distractors were closely comparable between groups, sleep dramatically increased the probability of falsely recognizing critical theme words (Table 1). Furthermore, the level of critical false recognition was practically identical to the level of true recognition under the wake condition and even clearly exceeded it under the sleep condition.

Insert Table 1 about here

To rule out the idea that sleep subjects' higher susceptibility to false memories was caused by unspecific circadian influences on learning and / or retrieval, in a supplementary experiment we assessed the levels of immediate learning on the same task at 08.30 h (n = 11) and 23.00 h (n = 11). The results (Table 2) demonstrate that neither true nor false recognition were affected by time of day. In particular, the proportions of falsely recognized critical lures were virtually identical under both conditions of time of day.

Insert Table 2 about here

Taken together, we have demonstrated that sleep has a strong facilitating influence on false recognition of non-presented theme words. On the other hand, subjects did not differ in their ability to distinguish studied from unstudied items when these were from different semantic categories (i.e. list words versus semantically unrelated distractors). Therefore, the false memory effect that we found cannot be explained by group differences in general recognition performance or response criterion (guessing tendency). Similarly, the results of the supplementary experiment clearly discourage the assumption of unspecific circadian effects acting on the process of false memory creation.

Beginning with early research on sleep and memory, sleep has been proven to substantially improve subjects' cognitive abilities. However, in the present study sleep caused a striking increase in memory intrusions. Can we therefore conclude to have found an ill effect of sleep? Our proposal is two-fold: Certainly sleep caused a deteriorated memory performance in terms of task demands which required subjects not to simply produce as many hits as possible but to discriminate between presented and non-presented words. On the other hand this impaired task performance may indicate optimized processes of abstraction acting during sleep. From this standpoint, sleep subjects outperformed their wake counterparts in grasping the gist of what they had studied. Evidence for this approach comes from patients suffering from semantic dementia. These patients have preserved recognition for objects and faces. Yet, they are selectively impaired in their semantic knowledge about the remembered items. In a recent study[23], patients with semantic dementia produced fewer false memories than their healthy controls, presumably due to their degraded capability of gist extraction. Still another promising attempt to understand the processes that underlie erroneous remembering comes from the spreading activation framework[24,25]. Following this conceptualization, semantic information is stored in an associative network with abundantly interconnected nodes, each representing a certain concept. As one semantic concept is being processed (e. g. sugar), activation spreads through the network thereby activating related concepts with a similar meaning (sweet). Interestingly, associative memory tests are among the classical tasks for which a sleep-induced enhancement could be demonstrated[8-11]. Thus, the false memory effect may actually be due to an augmentation of spreading activation during sleep.

Whatever the psychological and neuronal bases of false memory creation may be, we have demonstrated that the effect sleep has on memory is regardless of whether or not the events actually occurred. A recently published paper[20] raised the question: "Are false memories more difficult to forget than accurate memories?" (p. 1054). While the authors as well as others[21,22] came to a confirmative answer, we may now even go one step beyond: False memories are more persistent than accurate memories and they become even more so as you sleep on them.

Methods

In the main experiment a sample of 26 university students (7 female, 19 male, mean age: 25.2, s.d. = 2.5) was randomly divided into a sleep and a wake group of equal size. In the supplementary experiment another sample of 22 university students (13 female, 9 male, mean age: 23.7, s.d. = 2.3) was randomly allocated to the early (08.30 h) or late

(23.00 h) learning condition (11 subjects each). All subjects had a regular nightly sleep schedule and were free of psychiatric illness and psychoactive medication.

The memory test was the same for the main and the supplementary experiment. During learning, the 18 word lists[26] were spoken by a tape-recorded female voice with a delay of 750 ms between each word and with each list being separated by 10 sec break. For the recognition task, a computer program displayed the three words that had been presented at serial positions 1, 5, and 10 within one list during learning, the unrelated distractor words (two per list) as well as the critical lures (one per list) individually, in fixed random order. Thus, the recognition task contained 108 words, 54 of which had been studied and 54 of which had not (36 unrelated plus 18 critical distractors). For each word, subjects had to click one of two push buttons labelled “old” or “new” to indicate whether or not they had previously heard the word. There was an unlimited amount of time for each decision.

In the main experiment, subjects in the wake condition learned the word lists at 08.30 h and were tested for recognition at 16.00 h. Subjects in the sleep condition had their learning session at 23.30 h. They were put to bed at 00.00 h, immediately after they had finished learning. Sleep was polysomnographically monitored according to standard criteria[27] throughout the entire night and revealed good sleep quality (sleep onset latency, 12.85 ± 6.65 min; sleep efficiency, 90.06 ± 6.18%). Recognition was tested at 07.00 h. In the supplementary experiment, subjects studied the word lists at either 08.30 h or 23.00 h and performed the recognition test after three minutes of solving arithmetic tasks.

1. Stickgold, R., LaTanya, J. & Hobson, A. Visual discrimination learning requires sleep after training. *Nat. Neurosci.* **3**, 1237–1238 (2000).

2. Mednick, S. C. *et al.* The restorative effect of naps on perceptual deterioration. *Nat. Neurosci.* **5**, 677–681 (2002).

3. Fenn, K. M., Nusbaum, H. C. & Margollash, D. Consolidation during sleep of perceptual learning of spoken language. *Nature* **425**, 614–616 (2003).

4. Fischer, S., Hallschmid, M., Elsner, A. L. & Born, J. Sleep forms memory for finger skills. *Proc. Natl. Acad. Sci. USA* **99**, 11987–11991 (2002).

5. Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A. & Stickgold, R. Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron* **35**, 205–211 (2002).

6. Jenkins J. G. & Dallenbach, K. M. Obliviscence during sleep and waking. *Am. J. Psychol.* **35**, 605–612 (1924).

7. Benson, K. & Feinberg I. Sleep and memory: 8 and 24 hours after initial learning. *Psychophysiology* **12**, 192–195 (1975).

8. Benson, K. & Feinberg I. The beneficial effect of sleep in an extended Jenkins and Dallenbach paradigm. *Psychophysiology* **14**, 375–384 (1977).

9. Ekstrand, B. R. Effect of sleep on memory. *J. Exp. Psychol.* **75**, 64–72 (1967).

10. Ekstrand, B. R., Barrett, T. R., West, J. N. & Maier, W. G. in *Neurobiology of Sleep and Memory* (eds Drucker-Colin, R. R. & McGaugh, J. L.) 419–438 (Academic Press, New York, 1977).

11. Lovatt, D. J. & Warr, P. B. Recall after sleep. *Am. J. Psychol*. **81**, 253–257 (1968).

12. Wagner, U., Gais, S., Haider, H., Verleger, R. & Born, J. Sleep inspires insight. *Nature* **427**, 352–355 (2004).

13. Bartlett, C. F. *Remembering: A Study of Experimental and Social Psychology* (Cambridge University Press, Cambridge, England, 1932).

14. Spiro, R. J. Accommodative reconstruction in prose recall. *J. Verb. Learn. Verb. Be.* **19**, 84–95 (1980).

15. Sulin, R. A. & Dooling, D. J. Intrusion of a thematic idea in retention of prose. *J. Exp. Psychol.* **103**, 255–262 (1974).

16. Roediger, H. L., III. Memory illusions. *J. Mem. Lang.* **35**, 76–100 (1996).

17. Braun, K. A., Ellis, R. & Loftus, E. F. Make my memory: How advertising can change our memories of the past. *Psychol. Market.* **19**, 1–23 (2002).

18. Deese, J. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. Exp. Psychol*. **58**, 17–22 (1959).

19. Roediger, H. L., III. & McDermott, K. B. Creating false memories: Remembering words not presented in lists. *J. Exp. Psychol. Learn. Mem. Cogn.* **21**, 803–814 (1995).

20. Seamon, J. G. *et al.* Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Mem. Cognit.* **30**, 1054–1064 (2002).

21. Toglia, M. P., Neuschatz, J. S. & Goodwin, K. A. Recall accuracy and illusory memories: When more is less. *Memory* **7**, 233–256 (1999).

22. Payne, D. G., Elie, C. J., Blackwell, J. M. & Neuschatz, J. S. Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *J. Mem. Lang.* **35**, 261–285 (1996).

23. Simons, J. S. *et al*. Failing to get the gist: Reduced false recognition of semantic associates in semantic dementia. *Neuropsychology* **19**, 353–361 (2005).

24. Collins, M. A. & Loftus, E. F. A spreading-activation theory of semantic processing. *Psychol. Rev.* **82**, 407–428 (1975).

25. Roediger, H. L., III., Balota, D. A. & Watson, J. M. in *The Nature of Remembering* (eds Roediger, H. L., III., Nairne, J. S., Neath, I. & Surprenant, A. M.) 95–115 (American Psychological Association, Washington, DC, 2001).

26. Stadler, M. A., Roediger, H. L., III. & McDermott, K. B. Norms for word lists that create false memories. *Mem. Cognit.* **27**, 494–500 (1999).

27. Rechtschaffen, A. & Kales, A. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects.* (Public Health Service Publications, Washington, 1968).

Acknowledgements We thank Albert Lang for preparing audio-recordings of the word lists.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to O.L. (olaf.lahl@uni-duesseldorf.de).

Table 1: Effect of sleep on delayed recognition

Item type	Sleep	Wake	t_{24}	*P*
Studied				
List items	.65 ± .08	.62 ± .10	0.99	.33
Nonstudied				
Unrelated distractors	.30 ± .12	.27 ± .16	0.62	.54
Critical lures	.83 ± .09	.65 ± .11	4.70	< .0001***

Proportions of items classified as Old (mean ± s.d.) on the recognition test after a 7 h period of nocturnal sleep or diurnal wakefulness. *P* values are two-tailed.

Table 2: Effect of time of day on immediate recognition

Item type	Morning	Night	t_{20}	*P*
Studied				
List items	.71 ± .12	.77 ± .08	1.34	.20
Nonstudied				
Unrelated distractors	.14 ± .12	.15 ± .09	0.17	.87
Critical lures	.73 ± .20	.74 ± .16	0.13	.90

Proportions of items classified as Old (mean ± s.d.) on the immediate recognition test at morning (08.30 h) or at night (23.00 h). *P* values are two-tailed.

EQUIWORD: A software for the automatic creation of truly equivalent word lists

Olaf Lahl and Reinhard Pietrowsky

Institute of Experimental Psychology, University of Duesseldorf, Germany

Corresponding Author:

Olaf Lahl
Universität Düsseldorf
Klinische Psychologie
Universitätsstr. 1
40225 Düsseldorf
Germany
Phone: ++49-211-811-2146
Fax: ++49-211-8114261
E-mail: olaf.lahl@uni-duesseldorf.de

Abstract

Word lists are most commonly used in the investigation of human memory. To prevent transfer effects, repeated measures of memory for words require multiple lists of different words. Yet, the psycholinguistic properties of all word lists employed should match as closely as possible to avoid confounding with the independent variable(s) in question. Although comprehensive databases for word norms exist, to our knowledge no tool is available that automates the creation of such equivalent word lists. Instead, matching different lists is often accomplished prima facie. We have therefore developed a Windows program called EQUIWORD that completely automates the creation of word lists that are truly parallel with respect to a wide range of attributes. EQUIWORD takes psycholinguistic databases of different formats as input and computes several coefficients of distance for every possible word pairing. Program output consists of a list of all word pairs sorted according to their distance. On that basis, creating equivalent word lists is simply done by selecting the pairs with the lowest distance coefficients.

Introduction

Lists of single words are the most commonly used stimulus material when studying verbal memory. In a typical experimental setting, subjects learn a set of words during one or more study trials either by intention (explicit learning) or incidentally through a dummy task such as rating the words on their pleasantness or counting their letters (implicit learning). At the end of the retention interval (which may last only a few seconds or minutes in the case of immediate testing) memory performance is probed by explicit recall or recognition or by implicit tests, e.g. word-fragment completion.

Subject's ability to learn, retain, and retrieve a specific word largely depends on a variety of semantic and lexical features. Among others the most important characteristics are Paivio's scales of imagery and meaningfulness (Christian, Bickley, Tarka, & Clayton, 1978; Paivio, 1969, 1976), Osgood's scales of evaluation, potency, and activity (Bradley & Baddeley, 1990; Contini & Whissell, 1992; Eysenck, 1976; Howard-Voyer & Whissell, 1994; Lamarche, Campbell, Matheson, & Whissell, 1993; Levonian, 1972; Osgood & Suci, 1955) as well as word length (Lovatt & Avons, 2001) and frequency of use (Gregg, 1976). Even when these attributes are not part of the investigators' research object, there are at least two frequent occasions where the semantic and lexical *equivalence* of different word lists is essential. These are related to repeated measures as well as the sensitivity of measurements:

Due to possible transfer or carry-over effects (Hindmarch, 1980; Krauth, 2000), repeated measures of memory as part of experimental settings (e.g. Coen, Kinsella, Lambe, Kenny, & Darragh, 1990; Hirshman, Passannante, & Henzler, 1999; Peters & Levin, 1979; Plihal & Born, 1997; Soetens, Casaer, Hooge, & Hueting, 1995) or clinical and diagnostic testing (e.g. Crawford, Stewart, & Moore, 1989; Randt, Brown, & Osborne, 1980; Wechsler & Stone, 1973), require two or more lists containing different words. Yet, the words must match as closely as possible with respect to all psycholinguistic attributes that might affect recall, i.e. the lists must be parallel. Otherwise, with a large discrepancy within one or more critical variables, any differences between treatments cannot be attributed unambiguously to the experimental manipulation(s).

Even when only one word list is employed for a single measure of recall, a low intralist-variance of item difficulty is highly desirable to obtain sensitivity to the experimental manipulation(s). The sensitivity of a recall measurement might suffer when the individual items of a list differ considerably in one or more characteristics that govern how easily they may be learned or memorized. In that case, easy words will be remembered by almost all subjects and difficult words only by a few, regardless of the experimental treatments that were presumed to influence recall.

Although numerous pools of word norms for the relevant dimensions exist (Brown, 1984; Brown & Ure, 1969; Friendly, Franklin, Hoffman, & Rubin, 1982; Gilhooly & Logie, 1980; Kerr & Johnson, 1991; Kucera & Francis, 1967; Paivio, Yuille, & Madigan, 1968; Thorndike & Lorge, 1944; Toglia & Battig, 1978), manually parallelising words becomes a daunting and virtually infeasible task as soon as more than two variables are to be included. Or as Cutler (1981) points out: “At this point it is already clear to the experimenter that the task is probably impossible” (p. 68).

While this statement holds for handmade lists, the mathematical formulation of the problem is straightforward. A single word, measured for n different attributes, can be considered as a vector in an n-dimensional space. Hence, the overall dissimilarity between two words is the mathematical distance $d(\mathbf{x}_k, \mathbf{x}_l)$ of the two vectors $\mathbf{x}_k$ and $\mathbf{x}_l$ that represent them. Several coefficients for the distance of two vectors are commonly used in psychological applications and we will give an overview in the Appendix. For a list of m words there are $(m^2 - m) / 2$ pairs of different words. Once the pairwise distances are computed, the task of creating equivalent word lists is simply accomplished by picking up those pairs with the lowest (highest) coefficients of dissimilarity (resemblance). In addition, selection criteria are not restrained to preferably high proximities. It is also possible to create lists of words that have maximum distance or any other degree that is of interest for the researcher.

What we are providing here is a software program called EQUIWORD that accepts databases of word norms as input, computes the distance for each word pair and saves the sorted list of pairs to a file. EQUIWORD implements a set of five common distance measures and two methods of data standardization that are discussed in the Appendix. The program is capable of reading three different types of input files. The first file type adheres to the format used by the Medical Research Council (MRC) Psycholinguistic Database (Coltheart, 1981; Wilson, 1988). The MRC dictionary file is available on the Internet (http://ota.ahds.ac.uk/texts/1054.html) and contains a set of 150,837 English words measured for 26 psycholinguistic attributes. However, the entire set of attributes is not available for every word and EQUIWORD is restricted to the 14 attributes that are of quantitative nature. These attributes are: number of letters, number of phonemes, number of syllables, Kucera and Francis written frequency, Kucera and Francis number of categories, Kucera and Francis number of samples, Thorndike-Lorge frequency, Brown verbal frequency, familiarity, concreteness, imagery, mean Colorado meaningfulness, mean Paivio meaningfulness, and age of acquisition. For a detailed description of these psycholinguistic properties see Wilson (1988). EQUIWORD provides the possibility to include either the entire set or any possible subset of these attributes in the calculation of distance coefficients.

In addition to MRC database files, EQUIWORD accepts Microsoft Excel spreadsheet files and comma separated value files. With these file types the user can specify his or her own set of words and word attributes. As an example, Osgood's scales are not part of the MRC database. Yet, they are important for learning and remembering as outlined above. Provided that the user has access to ratings on these scales for a sample of words, he can easily create a database containing these ratings and use it with EQUIWORD. The usage of user-defined database files is described in detail in the following section.

Program Operation

After starting the program, MRC database files (file extension .dct), Microsoft Excel files[1] (.xls), or comma separated value files (.csv) can be loaded by selecting the "File Open" command. When working with an MRC database file, a subset of words can be selected by choosing one of four grammatical types – nouns, adjectives, verbs, and adverbs – and any combination of the 14 quantitative attributes the MRC database provides. Note, that for the computation of distance coefficients, EQUIWORD includes only the intersection of words with valid entries in all selected attributes.

Excel and comma separated value files offer the opportunity to load arbitrary collections of words (or even other objects) together with a set of ratings/attributes provided by the user. The first line of the file is the header and must specify the attributes' names. All following lines begin with a single word and contain its attribute values in the following columns. In the case of comma separated value files each attribute name in the header and each attribute value must be separated by a semicolon. See Figure 1 for two illustrative examples on how to format the data properly. Leaving an empty cell or a field between two semicolons indicates a missing value for that attribute. As with MRC database files, words with missing values in at least one attribute will be discarded in the computation of distance.

Insert Figure 1 about here

The Options menu offers choices for the type of data transformation (z-standardization, range standardization, or none), for the sort order of program output (sort ascending or descending by distance coefficient), as well as for the font type used for display. Finally, selecting one of the available distance coefficients from the Compute menu starts its computation. After finishing the calculation, the tabulated file output appears in the program window. Depending on the user's selection, all word pairs are listed in ascending or descending order based on the magnitude of their distance. Navigation to a specific word pair is possible through the Find Pair dialog that can be opened from the Edit menu. Selecting the Save command from the File menu saves the output to a text file for further processing by spreadsheet programs or text editors.

EQUIWORD also supports a second mode of operation that enables the user to compute the available distance coefficients all at once. To adopt simultaneous computation, one can select the option "All Coefficients at once" from the View Menu. Pressing the F5 key or selecting "All" from the Compute menu initiates calculation. The program output appears in a spreadsheet with one column following each coefficient. Clicking a column header initiates sorting the entire output by the corresponding coefficient in ascending order. Clicking the same column again toggles the sort order.

Discussion

With the software program presented in this paper, we hope to facilitate the creation of word lists that are truly equivalent in the mathematical sense of the word, thereby reducing potential sources of confounding in experiments that (have to) rely on psycholinguistic material. The value of our software for researchers is rising as more large word pools like the MRC database become available in electronic form and as an increasing number of psycholinguistic properties are included in the ratings. As an example, the inclusion of ratings for Osgood's scales of evaluation, potency, and activity (Osgood & Suci, 1955) is highly desirable. Ironically, these ratings are at least partially available for a pool of 650 words (Brown & Ure, 1969), but not in electronic form. We have already developed a test version of a software module for merging databases containing different word pools and ratings, and we are looking forward to implementing it in future versions of our program as soon as more word pools become available in a computer readable format.

System Requirements and Program Availability

EQUIWORD is a Windows program that requires the Microsoft .NET Framework 1.1, which can be freely downloaded from http://msdn.microsoft.com/netframework/downloads/framework1_1.

The EQUIWORD setup program checks the target machine for the presence of the .NET framework and redirects to this URL if it is not already installed. EQUIWORD itself requires about 4 MB of hard disk space. The program has been tested on a Pentium 4 processor with 1.7 GHz and 1 GB of RAM and we recommend using a similar amount of memory for handling large word sets. Memory consumption primarily varies with the number of words to be included in the computation. Therefore, if memory runs low when working with MRC database files, it is a good idea to restrict the number of words by selecting more attributes. Since only the intersection of words with valid entries in all selected attributes is included, this will lower the number of words and thereby the required amount of memory.[2]

EQUIWORD is free of charge for non-commercial use and can be downloaded via FTP from ftp://ftp.sleeplab.psycho.uni-duesseldorf.de/EquiWord or ftp://ftp.uni-duesseldorf.de/pub/psycho/lahl/EquiWord. Users planning to deploy EQUIWORD for commercial purposes (e.g. research in the psychopharmacological industry) are required to contact the first author. They will receive a CD-ROM containing the software for a charge of US $ 250,-. Authors who use EQUIWORD for their publication are expected to cite this reference and the program version used.

References

Bradley, B. P., & Baddeley, A. D. (1990). Emotional factors in forgetting. *Psychological Medicine*, **20**, 351-355.

Brown, G. D. A. (1984). A frequency count of 190,000 words in the London-Lund Corpus of English Conversation. *Behavior Research Methods, Instruments, & Computers*, **16**, 502-532.

Brown, W. P., & Ure, D. M. J. (1969). Five rated characteristics of 650 word association stimuli. *British Journal of Psychology*, **60**, 233-249.

Christian, J., Bickley, W., Tarka, M., & Clayton, K. (1978). Measures of free recall of 900 English nouns: Correlations with imagery, concreteness, meaningfulness, and frequency. *Memory & Cognition*, **6**, 379-390.

Coen, R. F., Kinsella, A., Lambe, R., Kenny, M., & Darragh, A. (1990). Creating equivalent word lists for the Buschke Selective Reminding Test. *Human Psychopharmacology*, **5**, 47-51.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, **33A**, 497-505.

Contini, L., & Whissell, C. (1992). Memory disadvantages for CVC associates of emotional words. *Perceptual & Motor Skills*, **75**, 427-431.

Crawford, J. R., Stewart, L. E., & Moore, J. W. (1989). Demonstration of savings on the AVLT and development of a parallel form. *Journal of Clinical & Experimental Neuropsychology*, **11**, 975-981.

Cutler, A. (1981). Making up materials is a confounded nuisance, or: Will we be able to run any psycholinguistic experiments at all in 1990. *Cognition*, **10**, 65-70.

Eysenck, W. M. (1976). Arousal, learning, and memory. *Psychological Bulletin*, **83**, 389-404.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The Toronto Word Pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, **14**, 375-399.

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, **12**, 395-427.

Gregg, V. (1976). Word frequency, recognition and recall. In J. Brown (Ed.), *Recall and recognition* (pp. 183-216). London: Wiley & Sons Ltd.

Hindmarch, I. (1980). Psychomotor function and psychoactive drugs. *British Journal of Clinical Pharmacology*, **10**, 189-209.

Hirshman, E., Passannante, A., & Henzler, A. (1999). The effect of midazolam on implicit memory tests. *Brain & Cognition*, **41**, 351-364.

Howard-Voyer, C., & Whissell, C. (1994). Emotional descriptors of words as predictors of recall in a paired-associate task. *Perceptual & Motor Skills,* **78**, 1187-1191.

Kerr, N. H., & Johnson, T. H. (1991). Word norms for blind and sighted subjects: Familiarity, concreteness, meaningfulness, imageability, imagery modality, and word associations. *Behavior Research Methods, Instruments, & Computers,* **23**, 461-485.

Krauth, J. (2000). *Experimental Design. A handbook and dictionary for medical and behavioral research*. Amsterdam: Elsevier.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lamarche, S., Campbell, M., Matheson, P., & Whissell, C. (1993). A memory advantage for serial lists composed of active words. *Perceptual & Motor Skills*, **77**, 748-750.

Levonian, E. (1972). Retention over time in relation to arousal during learning: An explanation of discrepant results. *Acta Psychologica*, **36**, 290-321.

Lovatt, P., & Avons, S. (2001). Re-evaluating the word length effect. In J. Andrade (Ed.), *Working memory in perspective* (pp. 199-218). Hove, U.K.: Psychology Press.

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, **2**, 49-55.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of Experimental Psychology*, **50**, 325-338.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Journal of Experimental Psychology*, **76**, 241-263.

Paivio, A. (1976). Imagery in recall and recognition. In J. Brown (Ed.), *Recall and recognition* (pp. 103-129). London: Wiley & Sons Ltd.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, **76**(1, Pt. 2).

Penrose, L. S. (1953). Distance, size and shape. *Annals of Eugenics*, **18**, 337-343.

Peters, B. H., & Levin, H. S. (1979). Effects of physostigmine and lecithin on memory in Alzheimer disease. *Annals of Neurology*, **6**, 219-221.

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

Randt, C. T., Brown, E. R., & Osborne, Jr., D. P. (1980). A memory test for longitudinal measurement of mild to moderate deficits. *Clinical Neuropsychology*, **2**, 184-194.

Romesburg, C. H. (1984). *Cluster analysis for researchers*. Belmont, CA: Lifetime Learning Publications.

Soetens, E., Casaer, S., Hooge, R. D., & Hueting, J. E. (1995). Effect of amphetamine on long-term retention of verbal material. *Psychopharmacology*, **119**, 155-162.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College Press, Columbia University.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.

Wechsler, D., & Stone, C. P. (1973). *Wechsler Memory Scale*. New York: Psychological Corporation.

Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, **20**, 6-10.

Author note

Correspondence concerning this article or obtaining the software should be addressed to Olaf Lahl, Heinrich-Heine-University, Department of Clinical Psychology, Universitätsstrasse 1 (Geb. 23.03), 40225 Düsseldorf, Germany (email: olaf.lahl@uni-duesseldorf.de).

Footnotes

[1] Loading Excel files requires the MS Excel application to be installed on the target machine. EQUIWORD checks this during program startup and offers the option of opening Excel files in the File Open dialog only if it finds an intact installation of MS Excel.

[2] On the other hand, 1 GB of RAM or more may be required if the number of possible word pairs is very large. This might happen on some rare occasions, such as when the entire MRC database (mrc2.dct) is used and the set of variables to be included for computation is available for a great number of words. An extreme example would be the matching of all nouns for their word length as the sole variable. The word length is of course available for the entire set of 22061 nouns, resulting in $(22061^2 – 22061) / 2 = 243.332.830$ word pairs. Since EQUIWORD requires 12 Bytes for the internal representation of each pair, this would result in a memory consumption of about 2.7 GB.

Figure Captions

Figure 1. Illustrative example on how to format user specific data files in **A** Microsoft Excel format or **B** comma separated value file format.

A

Noun	**Frequency**	**Imagery**	**Concreteness**	**Meaningfulness**	**Number of letters**
Umbrella	8	592	606	676	8
Reaction	124	395	312	484	8
…					

B

```
Noun;Frequency;Imagery;Concreteness;Meaningfulness;Number of letters

Umbrella;8;592;606;676;8

Reaction;124;395;312;484;8

...
```

Appendix

There are many ways to quantify the distance between objects, and the decision of which one to use depends mainly on the specific object of the research. EQUIWORD implements a set of five common measures that we will discuss briefly. For the following formulas we assume a data matrix **X** with words (objects) arranged in rows $i = 1,\ldots,m$ and attributes (object properties) arranged in columns $j = 1,\ldots,n$:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mj} & \cdots & x_{mn} \end{pmatrix} \quad (1)$$

Minkowski Metric

The general Minkowski- or L_r-Metric $d_r(\mathbf{x}_k, \mathbf{x}_l)$ is defined as

$$d_r(\mathbf{x}_k, \mathbf{x}_l) \equiv \sqrt[r]{\sum_{j=1}^{n} \left| x_{kj} - x_{lj} \right|^r} \quad (2)$$

where the Minkowski constant r is a positive integer. The special case of $r = 1$ is called the city-block metric:

$$d_1(\mathbf{x}_k, \mathbf{x}_l) = \sum_{j=1}^{n} \left| x_{kj} - x_{lj} \right| \tag{3}$$

The term city-block metric refers to the fact that this is the distance a pedestrian would have to cover on a grid of city streets that is confined to horizontal and vertical directions only. An airplane, on the other hand, would take the diagonal direction or air-line distance, which is the familiar Euclidian distance with $r = 2$. The Euclidian distance has a maximum size of the city-block distance (triangle inequality).

$$d_2(\mathbf{x}_k, \mathbf{x}_l) = \sqrt{\sum_{j=1}^{n} \left(x_{kj} - x_{lj} \right)^2} \tag{4}$$

The city-block metric applies the same weight to all differences. With increasing r, large differences contribute more to the magnitude of d_r which also means that the measure becomes more sensitive to outliers. In the case of r ? 8 , only the largest difference determines the distance. This is called the Chebychev distance:

$$d_\infty(\mathbf{x}_k, \mathbf{x}_l) = \max_j \left| x_{kj} - x_{lj} \right| \tag{5}$$

Mahalanobis Distance

The Euclidian distance tends to an overestimation of highly correlated variables. The Mahalanobis distance (Mahalanobis, 1936) is a generalization of the Euclidian distance that compensates for any of these inter-correlations by weighting the Euclidian distance with the inverse covariance matrix of the data $\mathbf{C}^{-1}$. It should therefore be applied in favour of the Euclidian distance when two or more variables of the data set are correlated to a considerable degree.

$$d_m(\mathbf{x}_k, \mathbf{x}_l) \equiv \sqrt{(\mathbf{x}_k - \mathbf{x}_l)^T \mathbf{C}^{-1} (\mathbf{x}_k - \mathbf{x}_l)} \tag{6}$$

In the case of uncorrelated variables, $\mathbf{C}$ and $\mathbf{C}^{-1}$ become diagonal matrices and the Mahalanobis distance is proportional to the Euclidian distance.

Coefficient of Shape Difference

The distance measures discussed so far will change their value if one vector is subjected to a contraction or dilation (multiplication by scalar *a* ? 1) or to an additive displacement (addition of vector **c** with all elements set to a constant *c* ? 0) (Romesburg, 1984):

$$d(\mathbf{x}_k, \mathbf{x}_l) \neq d(a\mathbf{x}_k + \mathbf{c}, \mathbf{x}_l) \tag{7}$$

with

$$\mathbf{c} = c\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{8}$$

In contrast, the coefficient of shape difference introduced by Penrose (1953) ignores additive displacements. It is defined as

$$\begin{aligned} d_p(\mathbf{x}_k, \mathbf{x}_l) &\equiv \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_{kj} - x_{lj})^2 - \left(\frac{1}{n}\sum_{j=1}^{n}(x_{kj} - x_{lj})\right)^2} \\ &= \sqrt{\frac{1}{n}d_2^2(\mathbf{x}_k, \mathbf{x}_l) - \left(\frac{1}{n}\sum_{j=1}^{n}(x_{kj} - x_{lj})\right)^2} \end{aligned} \tag{9}$$

From Equation (9) it can be seen that the first term is the squared average Euclidian distance. The second term denotes the squared average distance due to additive displacements and is subtracted from the first term. Consequently, two vectors that only differ by an additive constant c yield a Penrose distance of zero:

$$d_P(\mathbf{x}_k, \mathbf{x}_k + \mathbf{c}) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_{kj} - [x_{kj} + c])^2 - \left(\frac{1}{n}\sum_{j=1}^{n}(x_{kj} - [x_{kj} + c])\right)^2}$$

$$= \sqrt{\frac{1}{n}nc^2 - \left(-\frac{1}{n}nc\right)^2} = 0 \tag{10}$$

Cosine Coefficient

The cosine coefficient is a measure of resemblance. It gives the cosine of the angle *?* between two vectors:

$$d_c(\mathbf{x}_k, \mathbf{x}_l) \equiv \cos q = \frac{\mathbf{x}_k \cdot \mathbf{x}_l}{\|\mathbf{x}_k\|\|\mathbf{x}_l\|} = \frac{\sum_{j=1}^{n} x_{kj} x_{lj}}{\sqrt{\sum_{j=1}^{n} x_{kj}^2 \sum_{j=1}^{n} x_{lj}^2}} \tag{11}$$

thereby ranging from –1 (minimum similarity) to 1 (maximum similarity) and neglecting any differences in length (contractions or dilations due to multiplication by a scalar *a* > 0). A vector and its proportional transform will therefore obtain maximum similarity:

$$d_c(\mathbf{x}_k, a\,\mathbf{x}_k) = \frac{\mathbf{x}_k \cdot (a\,\mathbf{x}_k)}{\|\mathbf{x}_k\|\,\|a\,\mathbf{x}_k\|} = \frac{a\,(\mathbf{x}_k \cdot \mathbf{x}_k)}{|a|\,\|\mathbf{x}_k\|\,\|\mathbf{x}_k\|} = \frac{a\,\|\mathbf{x}_k\|^2}{|a|\,\|\mathbf{x}_k\|^2} = 1 \quad \text{for } a > 0$$

(12)

Correlation Coefficient

The familiar Pearson product-moment correlation is given by

$$d_R(\mathbf{x}_k, \mathbf{x}_l) \equiv r(\mathbf{x}_k, \mathbf{x}_l) \equiv \frac{\sum_{j=1}^{n} (x_{kj} - \bar{x}_{k\cdot})(x_{lj} - \bar{x}_{l\cdot})}{\sqrt{\sum_{j=1}^{n} (x_{kj} - \bar{x}_{k\cdot})^2 \sum_{j=1}^{n} (x_{lj} - \bar{x}_{l\cdot})^2}}$$

(13)

The correlation coefficient denotes the proportion of covariance of two vectors in relation to their total variance. Since any additive and proportional transformation of one vector affects both enumerator and denominator of Equation (13) likewise, the correlation coefficient is invariant to such transformations. A vector and its proportional and/or additive transform will therefore obtain maximum similarity:

$$d_R(\mathbf{x}_k, a\,\mathbf{x}_k + \mathbf{c}) = \frac{\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})(ax_{kj} + c - [a\overline{x}_{k\cdot} + c])}{\sqrt{\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})^2 \sum_{j=1}^{n}(ax_{kj} + c - [a\overline{x}_{k\cdot} + c])^2}}$$

$$= \frac{\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})(a[x_{kj} - \overline{x}_{k\cdot}])}{\sqrt{\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})^2 \sum_{j=1}^{n}(a[x_{kj} - \overline{x}_{k\cdot}])^2}}$$

$$= \frac{a\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})^2}{\sqrt{a^2\left(\sum_{j=1}^{n}(x_{kj} - \overline{x}_{k\cdot})^2\right)^2}} = 1 \quad \text{for } a > 0 \qquad (14)$$

Standardization of the Data Matrix

With the exception of the Mahalanobis distance, all coefficients discussed here are subject to bias if two or more attributes differ by their units of measurement. This is due to the fact that attributes with smaller units carry more weight in the calculation, and therefore EQUIWORD offers two methods of data normalization. The common z-standardization normalizes all attributes to the same unit "multiples of standard deviation" before performing any calculation. (Note that the Mahalanobis distance always does an implicit z-standardization.) Thus, the attribute values of each word are replaced by their corresponding z-score yielding the standardized data matrix **Z** with elements z_{ij} given by

$$z_{ij} \equiv \frac{x_{ij} - \bar{x}_{.j}}{s_{.j}} \tag{15}$$

On the other hand, the following standardization uses the minimum and maximum value of an attribute to confine all values to a range between 0.0 and 1.0, thereby expressing them in units of “proportion of range”:

$$p_{ij} \equiv \frac{x_{ij} - \min_{.j}}{\max_{.j} - \min_{.j}} \tag{16}$$

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfsmittel angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 24.04.2006

(Olaf Lahl)

Folgenden Personen möchte ich für Ihre Unterstützung danken:

Herrn Prof. Dr. Reinhard Pietrowsky für das angenehme Klima bei der Betreuung und das immer wache Interesse für mein Thema,

Frau Prof. Dr. Petra Stoerig für die angenehme Zusammenarbeit im Schlaflabor und die bereitwillige Übernahme der Zweit-Begutachtung,

Herrn PD Dr. Philipp Hammelstein und Herrn Dipl. Psych. Ronald Schneider für die vielen offenen und freundschaftlichen Fachgespräche,

meinen Diplomanden für Ihr großes Engagement und die vielen durchwachten Nächte im Labor.