

Relearn Faster and Retain Longer: Along With Practice, Sleep Makes Perfect

Stéphanie Mazza¹, Emilie Gerbier², Marie-Paule Gustin^{3,4},
Zumrut Kasikci¹, Olivier Koenig¹, Thomas C. Toppino⁵, and
Michel Magnin⁶

¹Laboratoire d'Étude des Mécanismes Cognitifs, Équipe d'Accueil 3082, Université Lyon 2, Université de Lyon;

²Bases, Corpus, Langage Lab, Department of Psychology, Unité Mixte de Recherche (UMR) 7320, Centre National de la Recherche Scientifique (CNRS), Université Nice Sophia Antipolis; ³Department of Public Health, Institut des Sciences Pharmaceutiques et Biologiques, Équipe d'Accueil 4173, Université de Lyon; ⁴Emerging Pathogens Laboratory – Fondation Mérieux, International Center for Infectious Diseases Research (CIRI), Institut National de la Santé et de la Recherche Médicale (INSERM) U111, UMR 5308, École Normale Supérieure de Lyon, Université Claude Bernard Lyon 1; ⁵Department of Psychology, Villanova University; and ⁶Central Integration of Pain (NeuroPain) Lab—Lyon Neuroscience Research Center, INSERM U1028, UMR 5292, CNRS, Université Claude Bernard Lyon 1

Psychological Science
1–10

© The Author(s) 2016

Reprints and permissions:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0956797616659930

pss.sagepub.com



Abstract

Both repeated practice and sleep improve long-term retention of information. The assumed common mechanism underlying these effects is memory reactivation, either on-line and effortful or off-line and effortless. In the study reported here, we investigated whether sleep-dependent memory consolidation could help to save practice time during relearning. During two sessions occurring 12 hr apart, 40 participants practiced foreign vocabulary until they reached a perfect level of performance. Half of them learned in the morning and relearned in the evening of a single day. The other half learned in the evening of one day, slept, and then relearned in the morning of the next day. Their retention was assessed 1 week later and 6 months later. We found that interleaving sleep between learning sessions not only reduced the amount of practice needed by half but also ensured much better long-term retention. Sleeping after learning is definitely a good strategy, but sleeping between two learning sessions is a better strategy.

Keywords

learning, sleep-wake cycle, relearning, sleep-dependent memory consolidation, repeated practice

Received 10/14/15; Revision accepted 6/27/16

As the old saying “practice makes perfect” suggests, extensive repetitive practice is often necessary for lasting retention of knowledge (Ebbinghaus, 1885/1913; Hintzman, 1976). However, in each relearning session, performance gains appear to become increasingly easy and fast, given that fewer practice trials are needed to reach an expected level of mastery. Such time saving during relearning (hereafter referred to simply as saving) represents a useful measure of retention dynamics (Bahrick, 1979; Ebbinghaus, 1885/1913; Nelson, 1985) and is supported by cerebral plasticity at both synaptic and system levels (Gerbier & Toppino, 2015; Lee, 2008, 2009). The schedule of the relearning sessions also has a considerable potential for improving long-term retention.

The spacing effect (Ebbinghaus, 1885/1913) illustrates that retention is boosted when practice sessions are spaced apart but diminished when sessions are massed in succession (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). The time that elapses between learning and relearning could account for this effect: Relearning is more difficult in the spaced condition. The *desirable difficulty* hypothesis assumes that the more effortful and difficult the learning, the deeper the cognitive processes

Corresponding Author:

Stephanie Mazza, Laboratoire d'Étude des Mécanismes Cognitifs, EA 3082, Université Lyon 2, Lyon, France
E-mail: stephanie.mazza@univ-lyon2.fr

engaged and the better the long-term retention (E. L. Bjork & Bjork, 2011; Schmidt & Bjork, 1992).

On the other hand, sleep has been shown to participate in consolidation and enhancement of new learning (Diekelmann & Born, 2010), benefiting memory through an off-line and effortless process (Rasch & Born, 2013; Stickgold, 2001). This sleep-dependent consolidation is thought to stem from the reactivation and integration of newly encoded memories into preexisting and permanent knowledge networks (Stickgold & Walker, 2013), a process that jointly involves system and synaptic-consolidation processes (Dudai, 2004). For instance, patterns of brain activity involved during learning are selectively replayed during subsequent sleep (Ji & Wilson, 2007; Maquet et al., 2000; Peigneux et al., 2004). This neuronal reactivation is associated with gradual structural redistribution of memories and changes in strength of synaptic connections within the hippocampo-neocortical network (Rasch & Born, 2013).

Thus, memory is enhanced both by repeated practice and by sleep, but the interaction between practice and sleep has not received much attention. Some early research suggested a beneficial effect of sleep on relearning (Benson & Feinberg, 1975; Castaldo, Krynicki, & Goldstein, 1974); however, the extent to which sleep, interleaved between practice sessions, contributes to time saving when relearning and favors long-term retention remains unknown. Bell, Kawadri, Simone, and Wiseheart (2014) investigated the benefit of sleep in a paradigm of spaced practice. Four schedules of learning and relearning were compared: massed (i.e., no interval), 12 hr during the same day (i.e., with no sleep), 12 hr overnight (i.e., with sleep), and 24 hr (i.e., with sleep). Long-term retention was assessed 10 days later. For participants who experienced the 12-hr overnight interval (which included sleep), long-term retention was increased compared with participants in the massed condition and was similar to that of participants who experienced the 24-hr interval. However, the authors failed to obtain the classic spacing effect in a comparison of the massed condition and the 12-hr interval filled with wakefulness. In addition, there was no difference in long-term performance after either of the 12-hr intervals (i.e., with or without sleep). Given these equivocal results, the authors concluded that neither sleep alone nor time between learning and relearning alone was sufficient to improve long-term memory.

To investigate more thoroughly the effect of sleep on saving during relearning, we used a successive relearning paradigm (see Bahrck, 1979) in which learning and relearning episodes were continued until a specified performance criterion was reached. The learning and relearning episodes were spaced by 12 hr of wakefulness or 12 hr that included a night of sleep. We hypothesized that the off-line and effortless sleep-dependent memory

consolidation would allow saving during relearning. We also wished to know how relearning after a period of sleep or wakefulness would translate into long-term retention. Recent research has shown that relearning to the point of mastery tended to attenuate (e.g., Rawson & Dunlosky, 2011, 2013) or even reverse (e.g., Storm et al., 2008) the initial benefit obtained by efficient methods of learning. In the present context, we explored whether relearning would weaken or, perhaps, amplify the benefits of previous sleep. Amplification of the benefits of sleep by relearning would suggest a critical role for sleep in learning efficiency, whereby sleep improves performance while reducing the effort needed to achieve a level of mastery.

Method

Overview

In the present study, participants were challenged to learn the French translation of 16 Swahili words using repeated retrieval-restudy practice. This practice has been demonstrated to be a particularly efficient learning method (Karpicke & Roediger, 2008). Twelve hours later, participants had the opportunity to relearn this material, and the amount of practice (i.e. number of trials) that participants needed to do so was measured. One group of participants was assigned to the wake condition and did not sleep between the initial learning session and the relearning session. Another group of participants was assigned to the sleep condition and did sleep between sessions. Later, a third group of participants was assigned to the control condition. One week later and 6 months after relearning, retention was further assessed for each group.

Participants

Forty healthy participants (18–29 years old) recruited at the University of Lyon participated in the main experiment. These participants were randomly assigned to one of the two groups (20 in each group). Later, 20 other participants (19–32 years old) were recruited as a control group. To ensure the homogeneity of the three groups, we tested the participants' sleep quality, circadian topology, level of sleepiness, and basic long-term and short-term memory capacity: Participants completed the Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989), the Horne and Ostberg morning/evening questionnaire (Horne & Ostberg, 1976), the Epworth Sleepiness Scale (Johns, 1991), and subtests from the Wechsler Adult Intelligence Scale IV (WAIS-IV; Wechsler, 2008) and the Wechsler Memory Scale III (WMS-III; Wechsler, 1997). Each participant provided

informed written consent. The study was conducted in accordance with the Declaration of Helsinki and was approved by the French ethics committee.

Groups

Participants in the wake group performed the initial learning session at 9:00 a.m. and the relearning session at 9:00 p.m. the same day. Participants in the sleep group performed the learning session at 9:00 p.m. and the relearning session at 9:00 a.m. the following day (i.e., after a night of sleep). Participants in the control group performed the learning session at 9:00 p.m., slept at night, and then performed a recall session at 9:00 a.m. the next morning.

Materials and procedure

Learning session. The initial learning session consisted of the presentation of the 16 Swahili-French word pairs (e.g., *nyanya-tomate*). Pairs were presented successively, in random order, each for 7 s, and participants were asked to study the French translation. After this presentation, each Swahili word was displayed with a response box under it, cuing the participants to type its French translation. Participants could take as long as they wished to answer or leave the response box empty. After either response, the correct pair (cue and target) was displayed for 4 s. The number of successful recalls (out of 16) during this first retrieval attempt was our measure of initial performance. The words that were correctly translated were not presented again, whereas the incorrectly translated and untranslated words were further practiced by the participants. Thus, they were trained to a criterion of correctly answering each of the 16 pairs by cycling through the list of items not yet answered correctly (i.e., once they provided a correct answer for a given pair of words, that pair no longer appeared). There was no limit on time or the number of retrieval attempts. The number of *pair trials* needed to complete the learning session was measured. There was a 12-hr interval between the learning and the relearning sessions. During this interval, actimetry (Actiwatch; CamNtech, Cambridge, United Kingdom) was used to quantify sleep duration in the sleep group.¹

Relearning session. On the first *list trial* of the relearning session, participants in the wake and sleep groups were asked to recall the French translation of each of the 16 Swahili words, and they received feedback (i.e., the correct answer) for each pair, in a procedure similar to that used in the learning session. Again, the number of successful recalls (out of 16) during the first retrieval attempt defined initial performance. However, in this relearning session, participants in the wake and sleep

groups practiced the whole 16-item list repeatedly until they successfully recalled the entire list. The number of list trials needed to attain this criterion was measured. As in the learning session, items were presented successively and randomly on each list trial. The learning and the relearning sessions were preceded by the assessment of subjective sleepiness level using the Karolinska Sleepiness Scale (Akerstedt & Gillberg, 1990). Participants in the control group had a similar recall session. However, it consisted of only one list trial, and there was no additional opportunity for further relearning.

Delayed testing. One week after and 6 months after the relearning session, all groups performed a midday cued-recall task without corrective feedback. The item order was random but the same for all participants.

Statistical analyses. For each dependent variable, normality was tested using the Shapiro-Wilk test, and variance homogeneity was tested using the modified Levene test. Nonparametric tests, such as Mann-Whitney U , rank-based analysis of variance (ANOVA; Conover & Iman, 1981), and r_s are reported when parametric tests could not be computed. For all statistical tests, p values of less than .05 were considered significant.

Results

Main experiment

The sleep and wake groups were not significantly different with regard to age (sleep group: $M = 22.30$ years, 95% confidence interval, or CI = [20.96, 23.64]; wake group: $M = 22.05$, 95% CI = [20.33, 23.77]), $t(38) = -0.24$, $p > .250$, or gender (sleep group: 10 women, 10 men; wake group: 9 women, 11 men), $\chi^2(1, N = 40) = 0.10$, $p > .250$. The two groups also did not differ with regard to sleep quality (Pittsburgh sleep quality index), circadian typology (Horne and Ostberg morning/evening questionnaire scores), or sleepiness (Epworth sleepiness scale)—sleep quality score (sleep group: $M = 3.40$, 95% CI = [2.78, 4.01]; wake group: $M = 3.85$, 95% CI = [3.09, 4.60]), $t(38) = 0.96$, $p > .250$; circadian typology score (sleep group: $M = 50.0$, 95% CI = [45.89, 54.10]; wake group: $M = 48.95$, 95% CI = [46.65, 51.24]), $U = 155$, $p = .222$; or sleepiness score (sleep group: $M = 6.75$, 95% CI = [5.24, 8.26]; wake group: $M = 8.35$, 95% CI = [6.76, 9.94]), $t(38) = 1.53$, $p = .135$. In addition, short- and long-term memory performance in the two groups was not statistically different on the Verbal Span subtest of the WAIS-IV (sleep group: $M = 9.63$, 95% CI = [8.95, 10.32]; wake group: $M = 10.0$, 95% CI = [9.13, 10.86]), $t(38) = 0.70$, $p > .250$. There was also no difference in the two groups' performance on the Verbal Paired Associates subtest of the WMS-III for immediate

recall (sleep group: $M = 29.79$, 95% CI = [28.44, 31.13]; wake group: $M = 30.16$, 95% CI = [29.11, 31.20]), $t(38) = 0.46$, $p > .250$, or for delayed recall (sleep group: $M = 7.89$, 95% CI = [7.74, 8.0]; wake group: $M = 8.0$, 95% CI = [8.0, 8.0]), $U = 161.5$, $p = .152$.

During the learning session, the sleep and wake groups did not differ significantly in the number of translations successfully recalled at the first retrieval attempt

after study (sleep group: $M = 4.55$ of 16, 95% CI = [3.42, 5.68]; wake group: $M = 3.80$ of 16, 95% CI = [2.55, 5.05]), $U = 152.5$, $p = .194$ (Fig. 1a). There was also no significant difference between the groups in the number of subsequent pair trials needed to answer each of the 16 pairs correctly (sleep group: $M = 5.55$, 95% CI = [4.62, 6.48]; wake group: $M = 5.15$, 95% CI = [4.45, 5.85]) $t(38) = 0.72$, $p > .250$ (Fig. 1b). Subjective sleepiness (Karolinska

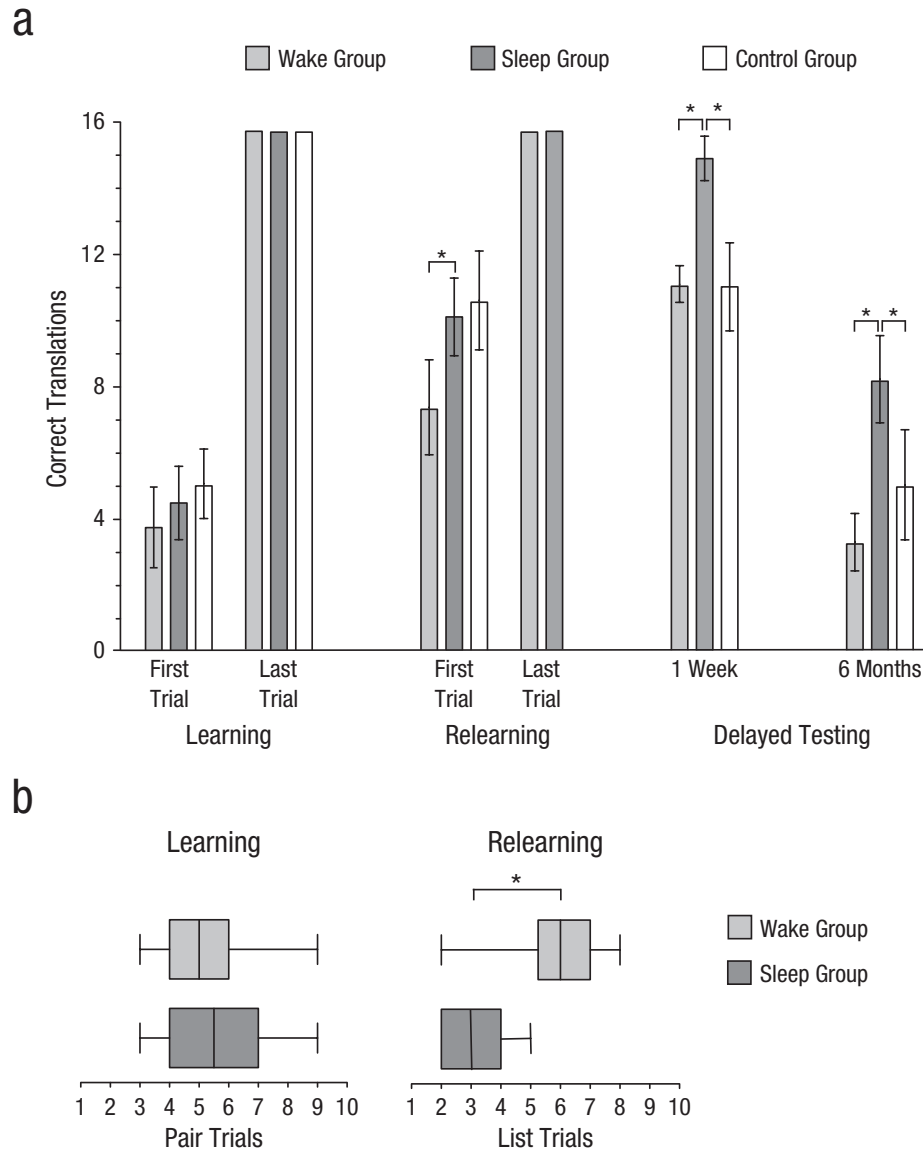


Fig. 1. Overall results. The graph in (a) shows the mean number of correct translations (out of 16 possible) during the first and the last practice trials in the learning session (pair trials) and relearning session (list trials) and during the cued-recall task after 1 week and 6 months. Results are presented separately for the wake, sleep, and control groups. The relearning session in the control experiment consisted of only the first list trial. Error bars represent 95% confidence intervals. The box-and-whiskers plots in (b) indicate the number of pair trials necessary for the wake group and the sleep group to attain the performance criterion in the learning session and the number of list trials necessary for them to attain the performance criterion in the relearning session. The left and right edges of the boxes represent the boundaries of the first and third quartiles, respectively, and the lines down the center of the boxes represent the medians. The left and right ends of the whiskers represent the minimum and maximum scores, respectively. Asterisks indicate significant differences between groups ($*p < .01$).

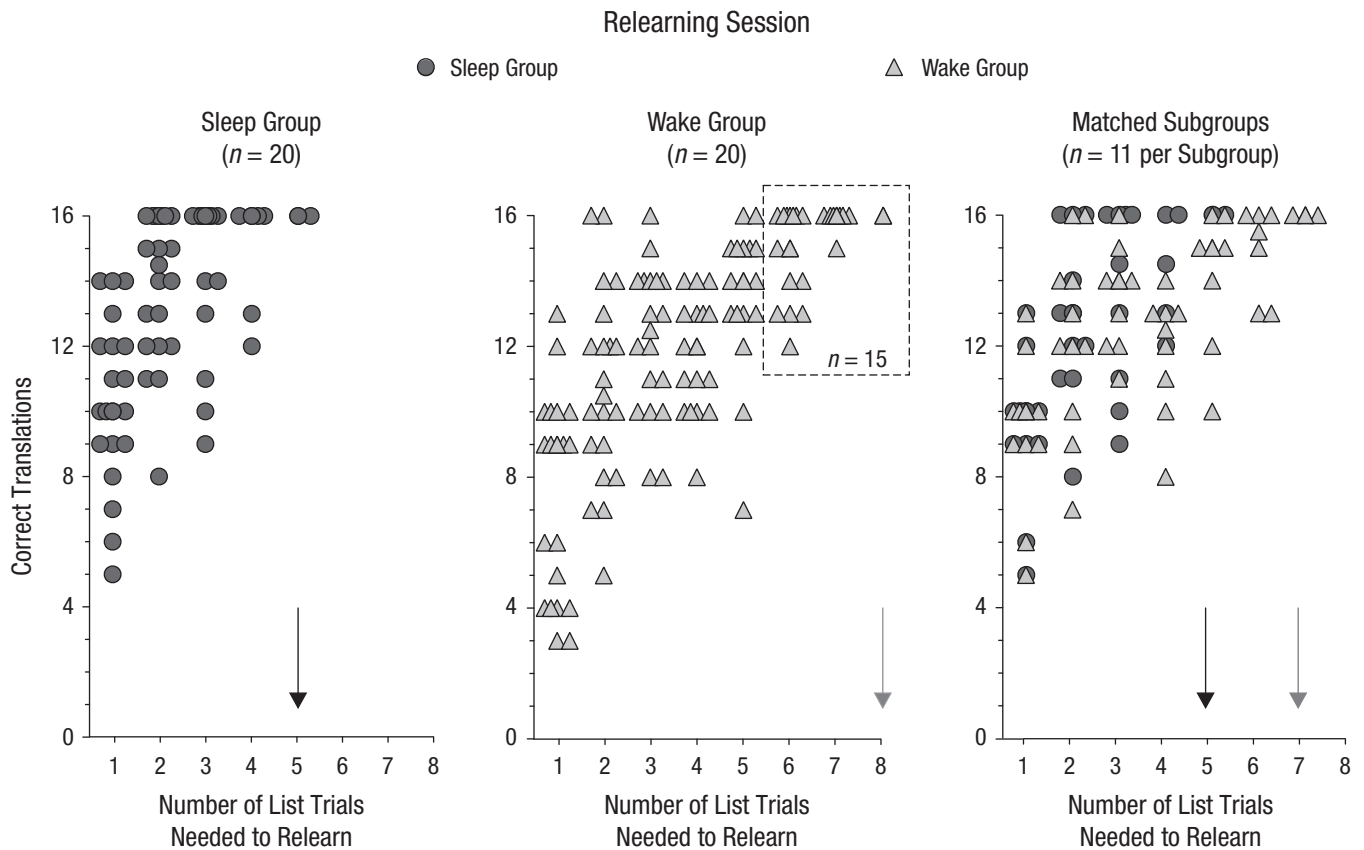


Fig. 2. Individual list-trial scores. The left and middle graphs show, respectively, the individual scores of members of the sleep and wake groups for each list trial in the relearning session. The maximum score was 16. The symbols enclosed in the dashed box indicate the successive scores for those participants in the wake group who still needed to continue after all of the participants in the sleep group had reached the criterion. The graph on the right shows individual scores of members of the sleep and wake subgroups for each list trial; the subgroups were matched on their performance in the first list trial. The arrows indicate the point at which all the participants in a given group reached the criterion.

Sleepiness Scale score) at the time of the learning session was similar in both groups, whether the session occurred in the evening for the sleep group ($M = 2.58$, 95% CI = [1.95, 3.21]) or in the morning for the wake group ($M = 2.84$, 95% CI = [2.23, 3.45]), $t(38) = 0.64$, $p > .250$. Thus, because there was no difference in circadian typology between the sleep and wake groups, these results are in agreement with previous findings (Bell et al., 2014; Payne, Chambers, & Kensinger, 2012), and our data are consistent with the assumption that the two groups completed the learning session with similar performance and effort and that performance was not significantly influenced by the time of day (evening or morning).

In contrast, after 12 hr, the number of initially remembered translations was higher for participants who had slept ($M = 10.3$, 95% CI = [9.09, 11.51]) than for those who had spent the day awake ($M = 7.45$, 95% CI = [5.98, 8.92]), $t(38) = 3.14$, $p = .003$, Cohen's $d = 0.99$ (Fig. 1a). Again, the level of sleepiness assessed at the beginning of this session was similar in both groups (sleep group: $M = 2.90$, 95% CI = [2.21, 3.57]; wake group: $M = 2.90$, 95%

CI = [2.36, 3.42]), $U = 175.5$, $p > .250$. This finding confirms the typical beneficial effect of postlearning sleep on retention (Diekelmann & Born, 2010) and is also consistent with the marginally significant group-by-session interaction obtained using a nonparametric rank-based ANOVA, $F(1, 38) = 3.97$, $p = .053$; this interaction suggests a larger difference between the groups at the beginning of the relearning session ($\Delta = 2.85$) compared with the beginning of the learning session ($\Delta = 0.75$).

Participants in the sleep group needed approximately half as many list trials to reach the relearning criterion (i.e., recalling all 16 translations; sleep group: $M = 3.05$, 95% CI = [2.58, 3.52]; wake group: $M = 5.80$, 95% CI = [5.02, 6.58]), $U = 44.5$, $p < .001$, $d = 2.00$ (Fig. 1b). Indeed, although all the participants in the sleep group completed the relearning session within a maximum of five list trials, 75% of the wake group needed further practice (Fig. 2). However, the sleep group may have been faster because fewer items had to be relearned. To address this issue, we compared a sleep subgroup ($n = 11$) and a wake subgroup ($n = 11$) matched with respect to their

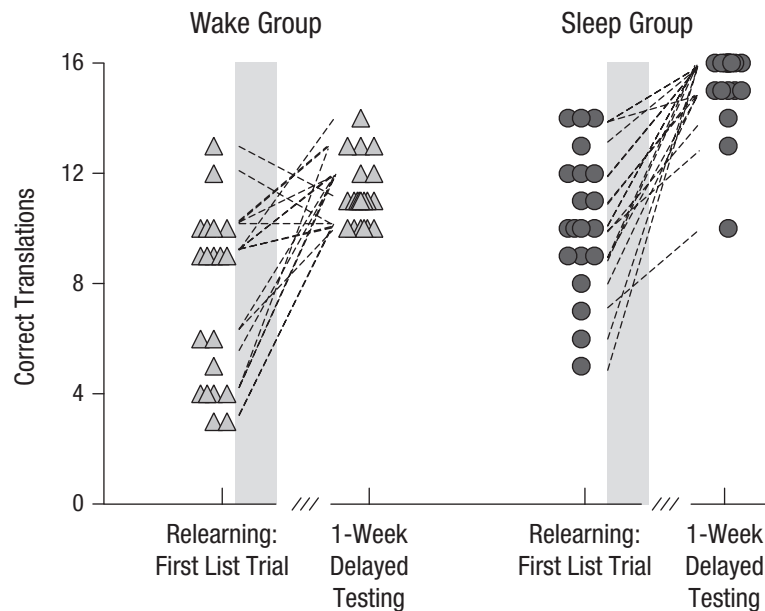


Fig. 3. Change in individual scores. Individual participants' number of correct translations on the first list trial of the relearning session and at the delayed testing at 1 week is graphed separately for the wake and the sleep groups. The gray shaded area in each graph represents the remaining list trials in the relearning session. The dashed lines connect the two scores for each participant.

initial performance at relearning.² The sleep subgroup still completed the relearning session more quickly ($M = 3.27$ list trials, 95% CI = [2.53, 4.01]) than did the wake subgroup ($M = 5.09$, 95% CI = [3.8, 6.38]), $t(20) = 2.72$, $p = .013$ (Fig. 2). The *relearning speed* (computed by dividing the number of unrecalled items at the first attempt by the number of list trials necessary to complete the session) was higher in the sleep subgroup ($M = 2.12$, 95% CI = [1.63, 2.60]) than in the wake subgroup ($M = 1.55$, 95% CI = [0.93, 2.18]), $U = 29$, $p = .037$. Thus, at the end of the relearning session, the wake group reached the same performance level as the sleep group (i.e., they reached the relearning criterion) but expended much more time and effort to do so.

One week later, the sleep group showed very little forgetting (15.20 correct translations, 95% CI = [14.51, 15.89]; Fig. 1a). Note that 60% of the participants performed perfectly at this delayed recall task. In sharp contrast, the wake group's retention suffered much more after the 1-week delay (11.25 correct translations, 95% CI = [10.68, 11.82]), $U = 20$, $p < .001$, $d = 2.93$; no participant was able to recall all 16 translations, and no score was higher than 14. Therefore, having relearned to criterion 1 week before was not sufficient for the wake group to reach the long-term performance of the sleep group. We were surprised that the difference between the groups at 1 week ($\Delta = 3.95$) appeared even greater than at the beginning of the relearning session ($\Delta = 2.85$). The

rank-based ANOVA revealed a marginally significant group-by-session interaction, $F(1, 38) = 3.55$, $p = .067$, which suggests a superadditive effect of sleeping and relearning. In addition, the ceiling effect at 1 week may have contributed to a lessening of the differences between the groups (Fig. 3).

Six months later, the sleep group ($n = 18$ as a result of attrition; 8.67 correct translations, 95% CI = [7.47, 9.86]) still outperformed the wake group (3.35 correct translations, 95% CI = [2.46, 4.24]), $t(36) = 7.61$, $p < .001$, $d = 2.46$ (Fig. 1a). Although the difference between the groups at 6 months ($\Delta = 5.32$) appeared larger than the difference at 1 week ($\Delta = 3.95$), the group-by-session interaction was not significant, rank-based ANOVA, $F(1, 36) < 1$, $p > .250$. However, in the sleep group, 56% of items recalled at 1 week were still recalled after 6 months, whereas in the wake group, only 30% of such items were recalled. The same pattern was observed even when the matched subgroups were compared at 1 week (sleep group: 15.45 correct translations, 95% CI = [14.76, 16.0]; wake group: 11.45 correct translations, 95% CI = [10.53, 12.38]), $U = 2.25$, $p < .001$, $d = 3.30$, and at 6 months (sleep group: 8.2 correct translations, 95% CI = [6.13, 10.27]; wake group: 3.73 correct translations, 95% CI = [2.31, 5.14]), $t(19) = 4.08$, $p < .001$, $d = 1.77$.

In the sleep group, the total sleep time (TST) interposed between the learning and the relearning sessions was positively correlated with relearning speed ($r = .47$,

$p = .038$) and with recall performance 1 week later ($r_s = .54, p = .015$; Fig. 4). Although one might wonder whether this correlation could be mediated by the participants' basic memory capacity, this possibility seems very unlikely because no significant association was identified between TST and the preexperimental assessments of short-term memory capability ($r = -.33, p = .171$) and long-term memory capability ($r = -.24, p > .250$).

Control experiment

To ensure not only that the advantage observed in the sleep group at 1 week was explained by the initial sleep-dependent consolidation but also that the relearning session performed after sleep had a specific influence on long-term retention, we conducted a control experiment with a third group of 20 participants (11 women and 9 men).

The control group and the sleep and wake groups were not significantly different with regard to age (control group: $M = 24.1$ years old, 95% CI = [22.68, 25.52]), $F(2, 57) = 2.42, p = .098$; circadian typology (control group: $M = 48.35$, 95% CI = [45.10, 51.60]), $F(2, 57) = 0.28, p > .250$; or level of sleepiness (control group: $M = 7.0$, 95% CI = [5.79, 8.20]), $F(2, 57) = 1.56, p = .219$. Short- and long-term memory performance was not statistically different on the Verbal Span subtest of the WAIS-IV (control group: $M = 9.7$, 95% CI = [8.8, 10.60]), $F(2, 57) = 0.25, p > .250$, or on the Verbal Paired Associates subtests from the WMS-III for immediate recall (control group: $M = 30.1$, 95% CI = [29.39, 30.81]), $F(2, 57) = 0.15, p > .250$, or for delayed recall (control group: $M = 8.0$, 95% CI = [8.0, 8.0]), $F(2, 57) = 2.18, p = .123$. There was a significant overall difference among the three groups for the Pittsburgh sleep quality index ($M = 2.8$, 95% CI = [2.40, 3.19]), $F(2, 57) = 3.28, p = .045$; there was a significant difference between the control group and the wake group, $t(38) = -2.56, p = .016$, but not between the control group and the sleep group, $t(38) = -1.72, p = .094$. Nevertheless, all participants could be considered "good" sleepers (see Buysse et al., 1989).

During the learning session, the control group's performance was not statistically different from that of the sleep and wake groups: mean number of successful initial translations = 5.15, 95% CI = [4.07, 6.23], $F(2, 57) = 1.50, p = .231, \eta^2 = 0.050$; mean number of subsequent pair trials needed to successfully answer each of the 16 pairs correctly = 5.0, 95% CI = [4.14, 5.86], $F(2, 57) = 0.51, p > .250, \eta^2 = .017$.

At the first list trial at recall 12 hr later, the three groups' performance was significantly different, $F(2, 57) = 7.09, p = .002, \eta^2 = .512$. The performance of the control group ($M = 10.75$, 95% CI = [9.22, 12.28]) was not significantly different from that of the sleep group, $t(38) = 0.47, p >$

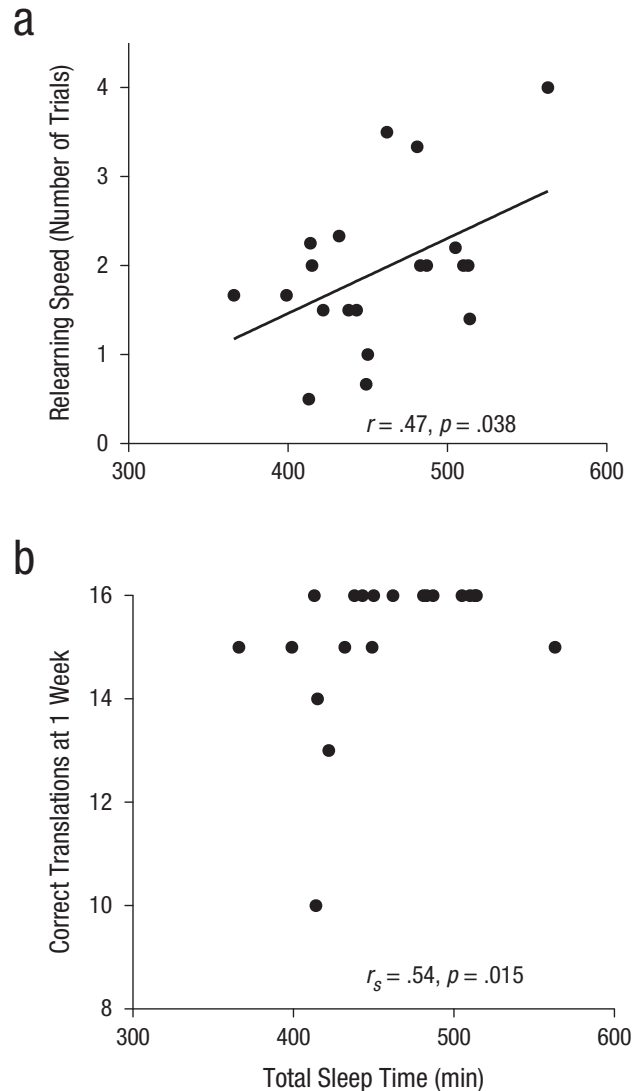


Fig. 4. Scatterplots showing the relationship between total sleep time and (a) relearning speed (with best-fitting regression line) and (b) recall performance at 1 week.

.250, but was significantly better than that of the wake group, $t(38) = 3.47, p = .002$, which confirms the beneficial effect of sleep compared with wakefulness.

One week later, however, the groups' performance was different, $F(2, 57) = 25.93, p < .001, \eta^2 = .476$. The control group's performance ($M = 11.25$, 95% CI = [9.89, 12.61]) was not statistically different from that of the wake group, $t(38) = 0.0, p > .250$, but was significantly lower than that of the sleep group, $t(38) = 6.24, p < .001$. This indicates not only that the sleep group's nearly perfect 1-week performance was the result of sleep-dependent consolidation but also that postsleep relearning made a specific contribution.

After 6 months, the groups' performance was again different, $F(2, 54) = 23.41, p < .001, \eta^2 = .464$. The control

group's performance ($M = 4.58$, 95% CI = [3.11, 6.04]) was not significantly different from that of the wake group, $t(37) = 1.52$, $p = .137$, but was significantly lower than that of the sleep group, $t(35) = 4.52$, $p < .001$. These results indicate that, for long-term retention, relearning intensively after 12 hr of wakefulness is no more productive than relearning very briefly after having slept.

Discussion

In previous work investigating the enhancing effect of sleep on subsequent declarative memory, the researchers focused on retention as reflected by recall or recognition (Diekelmann & Born, 2010; Rasch & Born, 2013). In the current study, the use of a relearning paradigm (Nelson, 1985) allowed us to obtain a more fine-grained measure of the effects of sleep on memory. Unrecallable items were reacquired faster during relearning in the sleep group independently of initial retention. In addition, both relearning speed and 1-week recall were positively correlated with time spent asleep. This shows that even memories not explicitly accessible at the beginning of relearning had also been transformed during sleep.

Our results underscore the specificity of the beneficial effect of sleep between learning and relearning for long-term retention. Bell et al. (2014), using a comparable design, found that neither sleep alone nor time between learning and relearning alone was sufficient to improve long-term memory. Their participants' high-level performance at the beginning of the relearning session may account for this absence of effect; such performance may have attenuated the impact of further relearning, especially because this relearning consisted of a fixed number of practice trials. Relearning is more than a sensitive method of testing memory: It is an integral part of learning and can be viewed as an iterative process in which we learn, forget, and then relearn as many times as necessary to achieve a specified level of retention. Our results indicate that when the interval between successive study sessions is filled with sleep rather than with wakeful activity, the process is much more efficient because it both facilitates relearning and enhances long-term retention.

Compared with the sleep group, the wake group started off the relearning session with a lower level of memory and therefore had to engage in greater retrieval efforts to reach the expected level of mastery. Those efforts, however, did not translate efficiently into long-term retention, as indicated by the fact that their performance did not reach the level attained by the sleep group 1 week later. Rather, the 1-week performance in the wake group was equivalent to that of the control group, which slept but then had only a single trial in which to relearn.

This result seems to be an exception to the *desirable-difficulty* principle (E. L. Bjork & Bjork, 2011; Schmidt & Bjork, 1992), according to which long-term retention is facilitated by conditions that make learning harder. In addition, relearning to criterion has been shown to attenuate the initial benefits induced by the use of efficient strategies at the time of learning, such as spaced learning (Rawson & Dunlosky, 2013) or test trials (Rawson & Dunlosky, 2011). Surprisingly, in the present study, the benefit of sleep was immune to any diminishing effects induced by relearning: Our results suggest even an amplification of the effect of postlearning sleep through relearning.

Sleep has been shown both to passively protect memories against decay and interference and to actively consolidate new memories (Diekelmann & Born, 2010). It is possible that the cognitive processes engaged by the wake group during relearning were inefficient because they operated on memories degraded by interference from diurnal activities (Wixted, 2004). In contrast, the lesser efforts expended by the sleep group may have been more rewarding in the long term because they operated on sleep-consolidated memories.

These results could also be interpreted within general memory models. For example, the new theory of disuse (R. A. Bjork & Bjork, 1992) assumes that two interacting parameters regulate memory performance. *Retrieval strength* determines the immediate accessibility of a memory and thus directly determines immediate performance. *Storage strength* is a latent parameter that reflects the degree of learning and moderates the effect of subsequent experience. All other things being equal, higher storage strength is associated with greater learning from subsequent practice and slower forgetting from subsequent interference; thus, higher storage strength yields greater memory stability over time. Because long-term retention is supported by the interplay between the two strengths, a high level of performance at the end of a training session—reflecting high retrieval strength—does not necessarily translate into long-term retention, which is determined by storage strength. We can speculate that sleep, through consolidation processes, increases the storage strength of memories. This would explain why the sleep group exhibited better recall before relearning than did the wake group, although less interference for the sleep group might also have contributed. Higher storage strength, however, also would explain the sleep group's better relearning efficiency, independently of initial performance. Finally, this would explain why, despite attaining the same relearning criterion, the two groups did not exhibit the same level of forgetting after 1 week and 6 months. How sleep affects these processes, and, more generally, how sleep interacts with learning strategies, clearly warrants further research.

Finally, the striking difference in long-term retention between the wake and sleep groups might also seem surprising given that the wake group slept shortly after relearning. However, in this group, sleep operated on memories that had suffered from more diurnal interference, the effects of which may not have been eliminated entirely by their extensive relearning. Moreover, participants in the wake group did not have the opportunity to relearn sleep-consolidated memories the next morning. These results suggest that an uninterrupted sequence of learning, sleep-dependent consolidation, and relearning (that is, repeatedly alternating study and sleep) is particularly efficient for long-term retention.

In conclusion, interleaving learning sessions with episodes of sleep may be an easy and promising method to achieve longer retention with less study. Such scheduling may be especially relevant for difficult or important to-be-learned declarative information. If comparable effects of sleep hold for procedural memory, it may be possible to enhance acquisition and retention of skills.

Action Editor

Colleen M. Kelley served as action editor for this article.

Author Contributions

S. Mazza, E. Gerbier, and O. Koenig designed the experiments. S. Mazza and Z. Kasikci performed all experiments. S. Mazza, E. Gerbier, M.-P. Gustin, and M. Magnin performed the analyses. S. Mazza, E. Gerbier, T. C. Toppino, and M. Magnin wrote the manuscript after discussion among all the authors. S. Mazza and M. Magnin supervised and coordinated the project.

Acknowledgments

We thank C. Bradley and G. A. Michael for comments on the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This study was supported by the LABEX CORTEX (Grant ANR-11-LABX-0042) of Université de Lyon, within the program "Investissements d'Avenir" (Grant ANR-11-IDEX-0007), operated by the French National Research Agency (ANR). S. Mazza was further supported by ANR Grant 15-CE33-0003-01.

Notes

1. The Actiwatch device uses an accelerometer to detect and log wrist movements. This method is known as *actigraphy* and has been shown to be a useful and sensitive means for discriminating sleeping from waking. Movements are sampled several times per second and are used to derive levels of activity and

inactivity, rhythm parameters (such as amplitude or acrophase), and sleep-wake parameters, such as the total sleep time. The total sleep time (TST) was determined as the number of minutes asleep in bed after lights go off, according to the sleep-wake categorization (Ancoli-Israel et al., 2003).

2. We selected from each of the two groups participants who displayed the exact same performance at the first list trial of the relearning session. For example, for a participant who had a score of 6 in the wake group, we selected a participant with a score of 6 in the sleep group. In order not to favor our hypothesis, if several participants met this criterion, we selected from the sleep group those for whom the number of required list trials was the highest and, from the wake group, those for whom this number was the lowest.

References

- Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29–37. doi:10.3109/00207459008994241
- Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. P. (2003). The role of actigraphy in the study of sleep and circadian rhythms. *Sleep*, *26*, 342–392.
- Bahrick, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, *108*, 296–308. doi:10.1037/0096-3445.108.3.296
- Bell, M. C., Kawadri, N., Simone, P. M., & Wiseheart, M. (2014). Long-term memory, sleep, and the spacing effect. *Memory*, *22*, 276–283. doi:10.1080/09658211.2013.778294
- Benson, K., & Feinberg, I. (1975). Sleep and memory: Retention 8 and 24 hours after initial learning. *Psychophysiology*, *12*, 192–195. doi:10.1111/j.1469-8986.1975.tb01275.x
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York, NY: Worth.
- Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*, 193–213. doi:10.1016/0165-1781(89)90047-4
- Castaldo, V., Krynicki, V., & Goldstein, J. (1974). Sleep stages and verbal memory. *Perceptual and Motor Skills*, *39*, 1023–1030. doi:10.2466/pms.1974.39.3.1023
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*, 354–380. doi:10.1037/0033-2909.132.3.354
- Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, *35*, 124–129. doi:10.2307/2683975

- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*, 114–126. doi:10.1038/nrn2762
- Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology*, *55*, 51–86. doi:10.1146/annurev.psych.55.090902.142050
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York, NY: Teachers College, Columbia University. (Original work published 1885)
- Gerbier, E., & Toppino, T. C. (2015). The effect of distributed practice: Neuroscience, cognition, and education. *Trends in Neuroscience & Education*, *4*, 49–59. doi:10.1016/j.tine.2015.01.001
- Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47–91). New York, NY: Academic Press.
- Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*, 97–110.
- Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, *10*, 100–107. doi:10.1038/nn1825
- Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. *Sleep*, *14*, 540–545.
- Karpicke, J. D., & Roediger, H. L., III. (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968. doi:10.1126/science.1152408
- Lee, J. L. C. (2008). Memory reconsolidation mediates the strengthening of memories by additional learning. *Nature Neuroscience*, *11*, 1264–1266. doi:10.1038/nn.2205
- Lee, J. L. C. (2009). Reconsolidation: Maintaining memory relevance. *Trends in Neurosciences*, *32*, 413–420. doi:10.1016/j.tins.2009.05.002
- Maquet, P., Laureys, S., Peigneux, P., Fuchs, S., Petiau, C., Phillips, C., . . . Cleeremans, A. (2000). Experience-dependent changes in cerebral activation during human REM sleep. *Nature Neuroscience*, *3*, 831–836. doi:10.1038/77744
- Nelson, T. O. (1985). Ebbinghaus's contribution to the measurement of retention: Savings during relearning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 472–479. doi:10.1037/0278-7393.11.3.472
- Payne, J. D., Chambers, A. M., & Kensinger, E. A. (2012). Sleep promotes lasting changes in selective memory for emotional scenes. *Frontiers in Integrative Neuroscience*, *6*, Article 108. doi:10.3389/fnint.2012.00108
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., . . . Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron*, *44*, 535–545. doi:10.1016/j.neuron.2004.10.007
- Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews*, *93*, 681–766. doi:10.1152/physrev.00032.2012
- Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General*, *140*, 283–302. doi:10.1037/a0023956
- Rawson, K. A., & Dunlosky, J. (2013). Relearning attenuates the benefits and costs of spacing. *Journal of Experimental Psychology: General*, *142*, 1113–1129. doi:10.1037/a0030498
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217. doi:10.1111/j.1467-9280.1992.tb00029.x
- Stickgold, R. (2001). Sleep, learning, and dreams: Off-line memory reprocessing. *Science*, *294*, 1052–1057. doi:10.1126/science.1063530
- Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience*, *16*, 139–145. doi:10.1038/nn.3303
- Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated relearning after retrieval-induced forgetting: The benefit of being forgotten. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 230–236. doi:10.1037/0278-7393.34.1.230
- Wechsler, D. (1997). *Wechsler Memory Scale—Third Edition*. San Antonio, TX: Psychological Corp.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition*. San Antonio, TX: Pearson.
- Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269. doi:10.1146/annurev.psych.55.090902.141555