Accelerated Long-Term Forgetting: Prolonged Delayed Recognition as Sensitive Measurement for Different Profiles of Long-Term Memory and Metacognitive Confidence in Stroke Patients

Nikki A. Lammers[1,2] , Selma Lugtmeijer[1,3] , Edward H.F. de Haan[1] and Roy P.C. Kessels[3,4,5,*]

[1]University of Amsterdam, Department of Brain and Cognition, Amsterdam, The Netherlands
[2]Amsterdam University Medical Center, Department of Neurology, Amsterdam, The Netherlands
[3]Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands
[4]Vincent van Gogh Institute for Psychiatry, Venray, the Netherlands
[5]Radboud Universit Medical Center, Department of Medical Psychology and Radboudumc Alzheimer Center, Nijmegen, the Netherlands

Abstract

Objective: Deficits in episodic memory are frequently reported after ischemic stroke. In standard clinical care, episodic memory is assessed after a 20–30 min delay, with abnormal memory decay over this period being characterized as rapid forgetting (RF). Previous studies have shown abnormal forgetting over a prolonged interval (days to weeks) despite normal acquisition, referred to as accelerated long-term forgetting (ALF). **Method:** We examined whether ALF is present in stroke patients ($N = 91$) using immediate testing (T1), testing after a short delay (20–30 min, T2), and testing after a prolonged delay (one week, T3). Based on performance compared to matched controls ($N = 85$), patients were divided into (1) patients without forgetting, (2) patients with RF between T1 and T2, and (3) patients with ALF at T3. Furthermore, confidence ratings were assessed. **Results:** ALF was present in a moderate amount of stroke patients (17%), but ALF was even more prevalent in our stroke sample than RF after a 20–30 min delay (which was found in only 13% of our patients). Patients reported a lower confidence for their responses, independent of their actual performance. **Conclusions:** Adding a one-week delayed measurement may potentially assist in identifying patients with memory decrements that may otherwise go undetected.

Keywords: Memory, Episodic, Stroke, Brain Ischemia, Memory Disorders, Humans, Mental Processes

INTRODUCTION

Deficits in episodic memory are among the frequent complications after ischemic stroke (Snaphaan & de Leeuw, 2007). Depending on the location, volume, and number of infarcts, encoding, consolidation,- and/or retrieval of verbal and visual information may be compromised (Lim & Alexander, 2009; Saczynski et al., 2009). Most episodic memory tasks rely on an acquisition phase, followed by an immediate recall test and/or recognition trial (T1), as well as a delayed recall test and/or recognition trial after approximately 20–30 min (T2) in order to assess long-term memory retrieval (Lezak, Howieson, Bigler & Tranel, 2012). Such a 20–30 min delay, however, may be too short to capture all deficits in long-term memory processes. This might partly explain why subjective memory complaints that are often reported after stroke are not always fully substantiated by performance measures, especially in the case of accelerated long-term forgetting (ALF; Butler & Zeman, 2008; Elliott, Isaac, & Muhlert, 2014; Geurts, van der Werf, Kwa, & Kessels, 2019).

The concept of ALF refers to an accelerated forgetting rate over longer delays (days to weeks) despite normal acquisition (Elliott et al., 2014). ALF is typically assessed by adding a distant third delayed test (T3) for recall/recognition. Individuals with ALF typically perform in the normal range on T1 and T2, but show significant impairments at T3 when compared to controls. Decreased performance, specifically at T3 (i.e., the presence of ALF), may reflect deficits in later stages of memory formation that are not captured with the standard delay of 20–30 min (Geurts, van der Werf, & Kessels, 2015). In the field of epilepsy, different time intervals, ranging from one day to six weeks, have been studied in

*Correspondence and reprint requests to: Roy P.C. Kessels, Radboud University, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognition – Neuropsychology & Rehabilitation Psychology, PO Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: r.kessels@donders.ru.nl

relation to ALF (Hoefeijzers, Dewar, Della Sala, Zeman, & Butler, 2013). The addition of such a long-delay measurement increased the sensitivity of memory tasks to identify patients with left temporal lobe epilepsy who experienced memory complaints in daily life (Blake, Wroe, Breen, & McCarthy, 2000). The patients performed in the unimpaired range on standard neuropsychological memory assessments, while showing significant increased long-term forgetting after an 8-week delay on a verbal memory task. Similarly, Fitzgerald, Thayer, Mohamed, and Miller (2013) found that the performance on a verbal memory task with extended delays of 24 h and 4 days was related to self-reported everyday memory complaints, whereas a 30-minute delay was not correlated with these complaints.

The underlying rationale is that during system consolidation, the circuitry that supports the memory reorganizes over time (Squire, Genzel, Wixted, & Morris, 2015). Different theories on memory consolidation diverge on their views on the reliance of memories on the hippocampus over time. During a 20–30-minute interval after initial encoding, consolidation may already take place (Dudai, 2012), but the time frame of the full consolidation process is unclear and might even be an ongoing process in which remote memories may continue to recruit hippocampal processing when retrieved (Dudai, 2012; Sekeres, Moscovitch, & Winocur, 2017). Since the full consolidation process might extend over such a protracted period of time, it is likely that the current default delay of 20–30 min does not cover all implicated memory processes.

There are currently two accounts for ALF compared to classic hippocampal amnesia, the latter being characterized by *rapid* forgetting (RF) (i.e., decay which is already present after a 20–30-minute delay, which for instance is observed in Alzheimer's dementia; Weintraub, Wicklund, & Salmon, 2012). The *qualitative* difference account suggests that ALF and classic hippocampal amnesia are two functionally distinct deficits, caused by different underlying mechanism (e.g. Fitzgerald et al., 2013). In this single dissociation, ALF patients show a similar memory performance at T2 compared to healthy individuals, but faster forgetting over longer delays. According to the *quantitative* difference account, ALF is forgetting that starts with subtle early retention deficits that progressively accelerate (Cassel & Kopelman, 2019; Cassel, Morris, Koutroumanidis, & Kopelman, 2016). This account states that classic hippocampal amnesia and ALF only differ with respect to their severity. It might be that the strength of the memory representation in ALF patients is already weaker at T2, but at that point yet too subtle to be detected.

An indirect proxy for the strength of a memory can be obtained from a measure of accuracy combined with a confidence rating (Migo, Mayes, & Montaldi, 2012). Higher confidence ratings are more often based on recollection or strong familiarity, while lower confidence ratings are almost entirely based on weaker familiarity (Migo et al., 2012). Although differences in memory confidence ratings are often related to accurate or inaccurate responding, several studies show that some participants tend to express more confidence than others, unrelated to their accuracy but dependent on subject-specific factors (e.g. Kantner & Dobbins, 2019; Kantner & Lindsay, 2014; Kelemen, Frost, & Weaver, 2000). A factor that is possibly related to confidence is the perception of one's own memory functioning. Geurts et al. (2019) demonstrated that TIA and minor stroke patients who performed normal on a verbal memory task after a short delay, but significantly worse than controls one week later, did not report more memory complaints than controls. In fact, patients who reported to be more content with their memory actually showed *more* forgetting over time. In contrast, healthy controls who were more content with their memory showed *less* forgetting over time. The authors suggest that healthy controls may be better at estimating their own memory functioning in comparison to the patients, but confidence ratings were not included in that study. Confidence ratings independent of accuracy give insight in subject-related factors, like perceived memory function, while the relation between accuracy and confidence provides information on memory strength.

So far, research into ALF in stroke patients is limited. The aim of the present study is twofold. One is to provide more insight into the prevalence of ALF after stroke. The second is to assess the relationship between memory performance and confidence ratings after different delays. To this end, a newly devised online set-up of a visual recognition task was administered in a large post-stroke patient group with supratentorial lesions. On the one hand, stroke may result in wide-spread disruptions of network activation, independent of the exact location of the lesion, on the other hand, stroke results in focal lesions that may be in a key region for episodic memory (Adhikari et al., 2017). Consequently, we hypothesize that episodic memory function can differ *between* stroke patients as a group and healthy controls, but also *within* stroke patients. Therefore, healthy participants were compared to stroke patients as one group, and to patients divided into three sub-groups: 1) patients without forgetting (no forgetting; NF), 2) patients with RF between T1 and T2, and 3) patients with ALF.

First, we hypothesize that a 1-week delayed measurement (T3) using a recognition task has added value for the detection of memory deficits that remain undetected with the standard 20–30-minute delay. Second, we expect that with increased delay duration confidence rates decrease, especially in patients with low memory performance.

MATERIALS AND METHODS

Study Design

Data collection was part of the multicentre cohort study "A functional Architecture of the Brain for Vision" (FAB4V) with the aim of assessing visual deficits in patients with ischemic stroke. Patients were admitted to one of the following hospitals in the Netherlands: University Medical Center Groningen (UMCG), Amsterdam University Medical Center (AmsterdamUMC), Radboud University Medical

Center (Radboudumc), University Medical Center Utrecht (UMCU), Onze Lieve Vrouwe Gasthuis (OLVG), Maasziekenhuis Pantein, Rijnstate, Ommelander Ziekenhuis Groep, and St. Antonius Ziekenhuis and Diaconessenhuis. The assessment took place at either the UMCG, AmsterdamUMC, Radboudumc, or the UMCU between September 2015 and February 2020. The study was approved by the Medical Ethics Review Committee Utrecht (METC-No. 2015.372).

Participants

Consecutive patients with a diagnosis of ischemic stroke were recruited for this study. Inclusion criteria were (1) the presence of a symptomatic cerebral (cortical or subcortical) ischemic stroke, diagnosed by a neurologist, (2) age between 18 and 90, and (3) fluent in Dutch. Exclusion criteria were (1) diagnosis of another neurological disease than an ischemic stroke, (2) diagnosis of a non-neurological disease that affects cognitive function, (3) history of substance abuse, (4) history of cognitive decline or impairment in daily life prior to the stroke (score ≥ 3.6 on the Dutch version of the Informant Questionnaire on Cognitive Decline in the Elderly [IQCODE], De Jonghe, Schmand, Ooms, & Ribbe, 1997), and (5) presence of severe disturbances in consciousness or comprehension to the extent that task instructions could not be understood. The assessment took place between two weeks and two years post-stroke.

A healthy control group, matched on age, was recruited without a history of any neurological diseases and psychiatric disorders that affect cognitive function and without a history of substance abuse. Written informed consents in accordance with the Declaration of Helsinki were obtained from all participants.

Neuroimaging

Stroke patients underwent a 3-T MRI scan as part of a large cohort study. At the Radboudumc and UMCG on the Siemens Magnetom Prisma, at the Amsterdam UMC and UMCU on the Philips R5. The sequence details for Siemens scanners were as follows: T2 FLAIR (TI = 1650ms, TR = 4800ms, TE = 484ms, [FOV] = 280mm, voxel size $0.9 \times 0.9 \times 0.9$mm^3). For the Philips scanners, sequence details were T2 FLAIR (TI = 1650ms, TR = 4800ms, TE = 253ms, [FOV] = 250mm, voxel size $1.12 \times 1.12 \times 0.56$mm3). Stroke lesions were delineated using ITK-SNAP software (Yushkevich et al., 2006). Using the clinical toolbox for SPM, the FLAIR and binary lesion mask were normalized to an older adult MNI template (Crinion et al., 2007; Rorden et al., 2012).

FLAIR images were obtained from 73 patients. Missing MRI data were due to time restrictions (10), technical problems (4), no clear lesion on the scan despite diagnosis of ischemic stroke (4).

Neuropsychological Assessment

As part of the cohort study, all patients underwent a short neuropsychological assessment. These include the Dutch version of the National Adult Reading Test as an estimate of premorbid IQ (Schmand, Lindeboom, & Van Harskamp, 1992), the WAIS-IV Digit Span Forward and Backward to assess working memory (Wechsler, 2008), a one-minute Category Fluency Test (animal naming; Van Der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006), the Trail Making Test for mental flexibility (Bowie & Harvey, 2006), and the Dutch 14-item Cognitive Screening Test (CST-14) (De Graaf & Deelman, 1991). The performance on these cognitive tests is included as an indication of the cognitive status of our patient sample.

Memory Assessment and Procedure

To assess ALF in visual episodic memory, we developed a computerized variant of the Doors Test (Schouten, Schiemanck, Brand, & Post, 2009), which is a subtest of the Doors and People Test (Baddeley, Emslie, & Nimmo-Smith, 2006). It consists of a four-alternative forced-choice paradigm, in which participants have to select a target picture of a door among three pictures of distractor doors (Figure 1). For our paradigm, we selected fifteen target doors and ninety distractor doors from the doors database of the University of York (https://www.york.ac.uk/res/doors). This database contains 2000 different doors, categorized on a range of variables including function, colour, age, condition, shape, door opening, glazing type, surrounding, and richness of detail. The fifteen target doors consecutively appeared in a random order at the center of a white computer screen. Each target door remained visible for 5 s. After the series of fifteen target doors, a series of fifteen four-alternative forced-choice arrays followed (T1). For each array, the patient had 13 s to select the target door from the distractors, by pressing 1, 2, 3, or 4 on the keyboard. In each array, the target door was matched with three distracter doors based on similarity in color, shape, and background details. After each response, participants indicated the confidence of their answer (1: "Very confident", 2: "Quite confident", 3: "Not confident" or 4: "It was a guess").

After a 20–30 min delay, participants were asked to identify the 15 target doors in a four-alternative forced-choice array including a new set of distractor doors (T2). After a mean delay of 7 days [range 5–13 days], all participants were called by telephone and asked to open a link to the online version of the delayed Doors test that was sent by email. They were asked to immediately complete this test. Again, participants had to identify the 15 target doors in a four-alternative forced-choice array (with a new set of distractor doors, T3). Participants were informed beforehand about the phone call, to ensure they were home and able to take the test. However, participants were not informed of the exact nature of the phone call (they were told we had some additional questions).
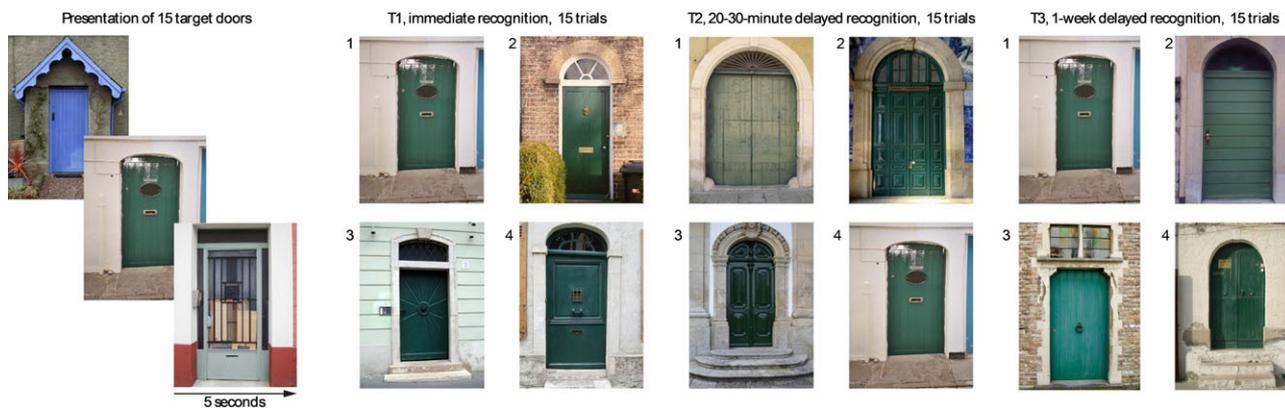
Figure 1. The Doors memory paradigm used in this study. On the left, the encoding samples only presented at T1, followed by one trial of the recognition memory task for each of the time measurements, in which the participant needs to identify the target among three distractors and give a confidence rating.

Statistical Analyses

Forgetting scores

In this study, we used forgetting scores, calculated relative to T1 as general baseline. The forgetting scores reflect the forgotten items as a percentage of the items that are acquired at T1 and are calculated with Equations 1 and 2.

$$\text{T2 forgetting score} = ((\text{hits at T1} - \text{hits at T2})/\text{hits at T1}) \times 100 \tag{1}$$

$$\text{T3 forgetting score} = ((\text{hits at T1} - \text{hits at T3})/\text{hits at T1}) \times 100 \tag{2}$$

The recognition task has a four-item forced-choice design; therefore, it might be that participants performed better at T2 or T3 compared to T1. As we are interested in forgetting after initial acquisition, negative forgetting scores were set to 0, indicating that 0% of the initially learned items are forgotten.

In order to compare the increase in the proportion of forgotten items from T1 to T2, and T1 to T3, we performed a mixed-model ANOVA with T2 forgetting score and T3 forgetting score as within-subject factor and group (all patients taken together versus controls) as between subject factor. Effect sizes were expressed as η_p^2. Bonferroni-corrected pairwise comparisons were used in case of a significant main effect of time. In case of violation of the assumption of sphericity Greenhouse–Geisser correction was applied. Additionally, to check whether forgetting was not affected by rehearsal that could be the result of the initial learning capacity (Baddeley, Atkinson, Kemp, & Allen, 2019), we conducted a linear regression analysis in order to assess the association between T1 accuracy (learning capacity) and forgetting rates between T2 andT3. Forgetting rates between T2 and T3 reflect the forgotten items at T3 as a percentage of the items that are acquired at T2 and are calculated with Equation 3.

$$\text{T2} - \text{T3 forgetting score} = ((\text{hits at T2} - \text{hits at T3})/\text{hits at T2}) \times 100 \tag{3}$$

Prevalence of RF and ALF within the patient group was calculated based on the performance of the stroke-free control group. A patient was identified with RF when he/she scored *below* 2 standard deviations of the mean of the control group on the T2 forgetting score. A patient was labeled as having ALF when he/she scored *within* 2 standard deviations from the mean of the control group on the T2 forgetting score and had a T3 forgetting score of *more than* 1.5 standard deviations above the mean of the T3 forgetting score of the control group (deviating forgetting score).

Comparability of the demographic characteristics of the patients without forgetting, patients with RF, patients with ALF, and healthy controls was assessed using one-way ANOVAs and chi-squared tests.

Confidence ratings

To establish the relative influences of memory accuracy, delay interval, and participant group (healthy controls, NF patients, RF patients, ALF patients) on memory confidence, we calculated the mean confidence for every participant for the trials with correct responses and those with incorrect responses. To assess how much participants adjust their judgment based on accuracy, a slope was calculated by subtracting the mean confidence for trials with correct responses from the mean for the trials with incorrect responses, per participant per time measurement.

Subsequently, two mixed-model ANOVAs were performed to test for the effect of group (healthy controls, NF patients, RF patients, ALF patients) and time (T1, T2, T3) on confidence ratings. Outcome measures were (1) the mean confidence rating, independent of accuracy and (2) the slope indicating the difference in confidence ratings between errors and hits. In case of significant main effects, Bonferroni-corrected pairwise comparisons were used to assess which moments of measurement or groups differed.

Imputation of missing data

Due to software complications, there was a small amount of missing data for hits and errors in our sample. A total of 25 values were missing. At T1, 7 values were missing divided over 5 trials. At T2, 12 values were missing divided over 7 trials. At T3, 5 values were missing divided over 3 trials. As the missing data were completely at random (MCAR), we used the multiple imputation method for imputation, with age, education, sex, group, and scores (hits/misses) on every single item as predictors.

RESULTS

Stroke Patients versus Healthy Controls

A total of 91 stroke patients and 85 age-matched healthy controls were included in the study. There were no statistically significant differences in forgetting rates between controls and patients when all patients were taken together [$F(1, 174) = .91$, $p = .341$, $\eta_p^2 = .005$]. A significant main effect of measurement time was found [$F(1, 174) = 68.5$, $p < .001$, $\eta_p^2 = .283$]. Both patients and controls had greater forgetting scores at T3 compared to T2. There was no time of measurement by group interaction effect [$F(1, 174) = .021$, $p = .885$, $\eta_p^2 = .000$]. T1 accuracy did not significantly predict the T2–T3 forgetting score in both controls [$R^2 = .000$, $F(1, 83) = .034$, $p = .855$] and patients [$R^2 = .001$, $F(1, 89) = .066$, $p = .798$]

Patient Sub-Groups

The prevalence of three different patterns of forgetting was examined. Totally, 15 patients (17%) showed ALF. They performed in the normal range on the T2 forgetting score, whereas their T3 forgetting scores were below 1.5 standard deviation from the mean of the control group. Twelve patients (13%) showed RF at T2. Sixty-three patients (70%) showed NF, as they had normal forgetting scores on both T2 and T3 compared to a healthy control group (see Table 1 for demographics of the patient subgroups and healthy controls and for cognitive profiles, mood states, and lesion characteristics of the patient subgroups). Figure 2 shows the lesion distribution of the three patient groups (see online supplementary materials S1 for lesion locations per individual in the RF and ALF subgroups). Figure 3 shows the mean accuracy of the patient subgroups and healthy controls per time of measurement.

Confidence Ratings

For mean confidence scores, independent of accuracy, a significant main effect of time of measurement was found [$F(1.62, 278.8) = 34.6$, $p < .001$, $\eta_p^2 = .167$]. Pairwise comparisons showed that mean confidence scores decreased from T1 to T2 ($p = .002$), from T1 to T3 ($p < .001$), and from T2 to T3 ($p < .001$). A significant main effect for group was found [$F(3, 172) = 7.1$, $p < .001$, $\eta_p^2 = .110$]. Healthy controls were

more confident than NF patients ($p = .002$), RF patients ($p = .023$), and ALF patients ($p = .027$). No significant differences in confident ratings were found between the three subtypes of patients. No significant time by group interaction effect was found [$F(1, 174) = .02$, $p = .885$, $\eta_p^2 = .000$].

Analysis of the slopes (see Figure 4), indicating the difference in confidence per correct versus incorrect responses, showed a significant main effect of time of measurement [$F(2, 304) = 6.64$, $p = .002$, $\eta_p^2 = .283$]. Pairwise comparisons showed that the slopes at T3 significantly differed from those at T1 ($p = .010$) and T2 ($p = .003$), whereas there was no difference between those at T1 and T2. At T3, the difference in confidence scores between correct and incorrect responses was smaller than at T1 and T2. Neither a main effect of group [$F(3, 152) = 1.77$, $p = .156$, $\eta_p^2 = .034$] nor an interaction effect of time of measurement by group was found [$F(6, 304) = .56$, $p = .764$, $\eta_p^2 = .011$].

DISCUSSION

In this study, we evaluated the prevalence of ALF after stroke in comparison to the prevalence of RF and normal forgetting (relative to stroke-free controls). Furthermore, we investigated the effect of a longer delay on meta-cognitive confidence for these participants groups, using a visual recognition memory paradigm.

We demonstrated that ALF is present in a moderate amount of stroke patients. With 17% of the patients showing abnormal forgetting after a one-week delay without showing deviating scores on the standard delay, ALF appears to be even more prevalent in our stroke sample than RF shown already after a 20–30 min delay (which was the case in 13% of the patients). Adding a one-week delayed measurement may thus potentially be valuable for the identification of patients with memory decrement that are missed in standard clinical practice where a 20–30 min delay is the default. At group level, we did not find any significant differences in forgetting rates between stroke patients and controls. This is in line with previous findings on verbal recognition (as opposed to recall) tasks with a prolonged delay, where no difference was found between controls and patients with TIA or minor stroke (Geurts et al., 2019).

Our data support the notion by Fitzgerald et al. (2013) that ALF and RF are two distinct phenomena. This in contrast to the view of ALF as differing from RF only with respect to initial (i.e. at T2) severity of forgetting, as argued by Cassel et al. (2016) and Cassel and Kopelman (2019). One should note, however, that our study is not directly comparable with the work by Cassel et al. (2016) and Cassel and Kopelman (2019), as the studies differ on several methodological aspects, such as type of memory task, number of intervals, and matching procedure. However, in our sample, patients with ALF and RF show very different patterns of forgetting. Stroke patients who demonstrate ALF showed similar T2 forgetting rates as stroke-free individuals and stroke patients without memory problems, but high forgetting rates

Table 1. Demographical characteristics of the no-forgetting patient group (NF), the patients with rapid forgetting (RF), the patients with accelerated long-term forgetting (ALF), and the healthy controls, and lesion characteristics and neuropsychological tests scores of the patient sub-groups

	NF patients ($N = 64$)	RF patients ($N = 12$)	ALF patients ($N = 15$)	Controls ($N = 85$)	p value
Age (mean $\pm$ SD)	62.3 $\pm$ 10.8	60.8 $\pm$ 6.4	66.9 $\pm$ 11.7	62.0 $\pm$ 11.4	.399
Sex (m/f)	43/21	11/1	9/6	42/43	.016
Education level N					.328
Less than primary school	0	0	1	0	
Primary education completed	1	1	1	0	
More than primary education (no diploma)	1	1	0	1	
Lower secondary education completed	9	2	4	6	
Average-level education completed	13	2	4	24	
High-level secondary education completed	28	1	1	37	
Academic degree	12	5	4	17	
Years of education (mean $\pm$ SD)	14 $\pm$ 3.3	13 $\pm$ 4.7	12 $\pm$ 4.5	14 $\pm$ 3.1	.183
Interval stroke-assessment in weeks (median [range])	10.0 [3.0–109.9]	9.6 [4.1–110.1]	9.0 [4.0–112.9]	-	.886
Lesion volume mm^3 (median [range])	5,3875 [69–87,251]	7,344 [409–137,487]	3,7436 [2–85,122]	-	.601
Location infarction N (%)					
Left hemisphere	26 (40.1)	3 (25.0)	3 (20.0)	-	
Right hemisphere	16 (25.0)	5 (41.7)	5 (33.3)	-	
Bilateral	12 (18.8)	1 (8.3)	2 (13.3)	-	
No MRI available	10 (15.6)	3 (25.0)	5 (33.3)	-	
HADS Depression scale (N, $M \pm SD$)	44, 3.3 $\pm$ 2.8	6, 6.2 $\pm$ 4.5	8, 2.9 $\pm$ 2.7	-	.081
HADS Anxiety scale (N, $M \pm SD$)	44, 4.4 $\pm$ 3.7	6, 3.2 $\pm$ 2.2	8, 5.5 $\pm$ 2.5	-	.459
NART-IQ (N, $M \pm SD$)	60, 108.6 $\pm$ 13.3	10, 104.0 $\pm$ 15.1	14, 103.7 $\pm$ 16.9	-	.389
Digit Span Forward (N, $M \pm SD$)	59, 8.4 $\pm$ 2.1	8.4 $\pm$ 1.5	13, 7.4 $\pm$ 1.5	-	.242
Digit Span Backward (N, $M \pm SD$)	59, 7.6 $\pm$ 2.0	10, 7.0 $\pm$ 2.2	13, 7.0 $\pm$ 1.4	-	.450
Semantic Fluency (N, $M \pm SD$)	58, 21.3 $\pm$ 5.4	9, 21.7 $\pm$ 3.3	11, 20.5 $\pm$ 4.3	-	.826
Trail Making Test - part A (N, $M \pm SD$)	58, 37.7 $\pm$ 13.5	11, 45.8 $\pm$ 28.2	13, 49.5 $\pm$ 19.1	-	.048
Trail Making Test - part B (N, $M \pm SD$)	58, 91.8 $\pm$ 38.9	11, 99.9 $\pm$ 75.7	13, 118.4 $\pm$ 37.3		.161
Cognitive Screening Test (N, $M \pm SD$)	49, 13.7 $\pm$ .7	8, 13.6 $\pm$ 1.1	12, 13.7 $\pm$ .4	-	.820

Education level is expressed as 7 categories, based on the Dutch educational system (Duits & Kessels, 2014) that is comparable with the International Standard Classification of Education (UNESCO, 2011). The estimated years of education for comparison with the Anglo-Saxon educational system are presented for descriptive purposes (cf. Hohstenbach et al., 1998); HADS: Hospital Anxiety and Depression Scale (Zigmund & Snaith, 1983); NART: National Adult Reading Test (Dutch version; Schmand, Lindeboom, & Van Harskamp, 1992).

during the prolonged delay. In contrast, stroke patients with RF did not show any further decline after T2. Although it was expected that ALF patients did not significantly differ from patients without memory problems on T2 forgetting scores (as we defined both groups by the fact that they performed *within* 2 standard deviations from the mean of the control group), according to the quantitative theory one would expect that the strength of the memory representation in ALF patients is already weaker at T2, though very subtly (Mayes, Hunkin, Isaac, & Muhlert, 2019). There were no significant differences between the patient subgroups on any of the other cognitive or mood measures. Although the RF group's mean absolute depressive symptoms score is slightly higher than the other group means, this difference was statistically not significant and none of the patients met the criteria for major depressive disorder. This indicates that the different patterns of forgetting do not reflect general cognitive

dysfunction, differences in mood, or differences in time since stroke or lesion volume.

From a neural perspective, a qualitative difference between ALF and RF is also plausible. That is, RF has been associated with medial temporal lobe functioning (Squire et al., 2015), while ALF, as a deficit in the later stage of memory consolidation, might rely more on distributed cortical networks. Previous studies on epilepsy suggest that ALF is not related to hippocampal (Butler et al., 2009; Wilkinson et al., 2012) or temporal-lobe volume (Wilkinson et al., 2012), which is in line with our finding that ALF may occur in stroke patients without hippocampal lesions. Our study could not identify specific lesion characteristics (i.e., location or size) that explain which patients show which pattern of forgetting. One of the theories on ALF in epileptic patients is functional disconnection. This hypothesis suggests that seizures disrupt the neurophysiologic processes
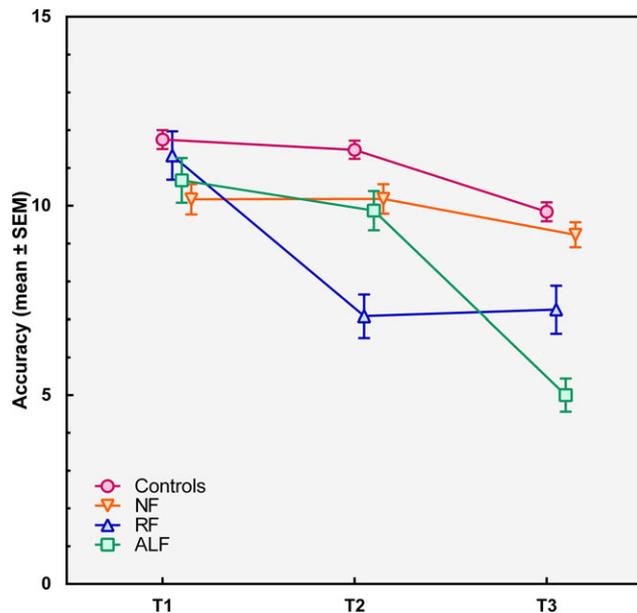
Figure 2. Lesion overlap map per subgroup, left is depicted on the left, coordinates are in MNI space, a) no-forgetting patient group (NF), b) the patients with rapid forgetting (RF), the patients with accelerated long-term forgetting (ALF).
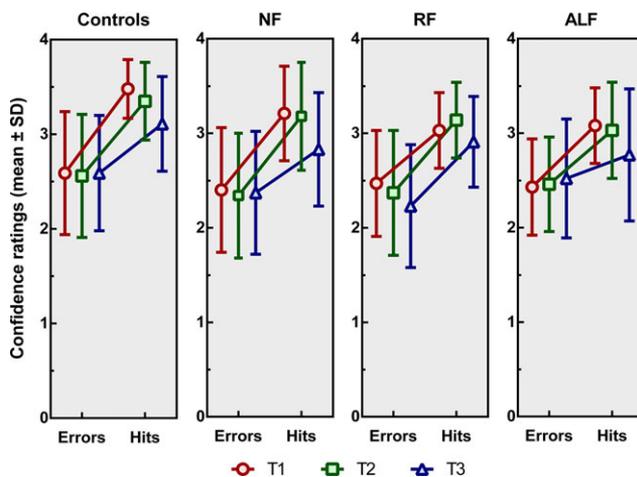

Figure 3. Recognition accuracy (means ± standard errors of the mean) immediately after the encoding phase (T1), after 20–30 min (T2) and after one week (T3) for the no-forgetting patient group (NF), the patients with rapid forgetting (RF), the patients with accelerated long-term forgetting (ALF), and the healthy controls.

underlying long-term memory consolidation. Fitzgerald et al. (2013) demonstrated that subclinical epileptiform discharges disrupt the consolidation of memory, resulting in ALF. In stroke patients, reduced functional connectivity in the default mode network has been associated with episodic memory dysfunction, as measured by the delayed recall score on California Verbal Learning Test (Tuladhar et al., 2013). Future research should investigate whether RF and ALF

are associated with different patterns of reduced functional connectivity after stroke.

Regarding meta-cognitive confidence after stroke, stroke patients rate their memory confidence overall lower than healthy controls, regardless of their actual performance. An explanation for this finding may be that psychological and emotional effects of stroke may affect one's self-appraisal and confidence (Wei et al., 2016). Furthermore, in all participants (healthy controls and the three groups of stroke patients), confidence decreased when the delay after encoding increased. Visual inspection of the data suggests that ALF patients rated their own memory performance lower compared to the other subtypes of patients and healthy controls, although this finding was not statistically significant. Especially at the one-week delayed test, ALF patients seemed less confident when giving correct answers compared to the other participants. This might indicate that the correct responses of ALF patients are more often based on weak familiarity, while the correct responses of the other patients might be more often based on recollection or strong familiarity. As familiarity is an indirect proxy for weaker memory capacity (Migo et al., 2012), this finding complements our previous finding in such that memory of ALF patients is not only less accurate at T3 compared to other patients but is also weaker. This, however, needs to be replicated and examined in more detail in future research.

There are some limitations and strengths in the design of this study. A practical issue concerning recognition tasks in long-term forgetting is the effect of rehearsal that could be a result of (residual) learning capacity (Baddeley, Atkinson, Kemp, & Allen, 2019). However, as the aim of our research was to detect ALF, multiple tests for the same target items within the same person were necessary. Generally, this could result in a bias that patients with impaired learning capacity but normal rates of forgetting are easily mistaken for ALF (Baddeley et al., 2019). However, our data do not show faster forgetting in poorer learning performers. Hits on T1 (learning capacity) did not predict the forgetting rates between T2 and T3 in both the patient group and the healthy control group. Additionally, our data do not indicate a learning affect, as the control group did not show any increase in accuracy between T1, T2, and T3.

A second issue in memory research is that recognition tests are often less effortful and more susceptible for ceiling effects than recall tests, which could reduce the sensitivity of ALF measures (Elliott et al., 2014). Exploration of the distribution of the scores on T1, T2, and T3 in our sample, however, did not reveal any ceiling effects for any of the three time measurements. Furthermore, it is not possible to examine visual memory for stimuli such as the ones we used using a free-recall test.

A final important practical issue in ALF research concerns the debate on how forgetting rates can be compared when learning capacity may differ between groups (Loftus, 1985; Wheeler, Ewers, & Buonanno, 2003). As a solution for this
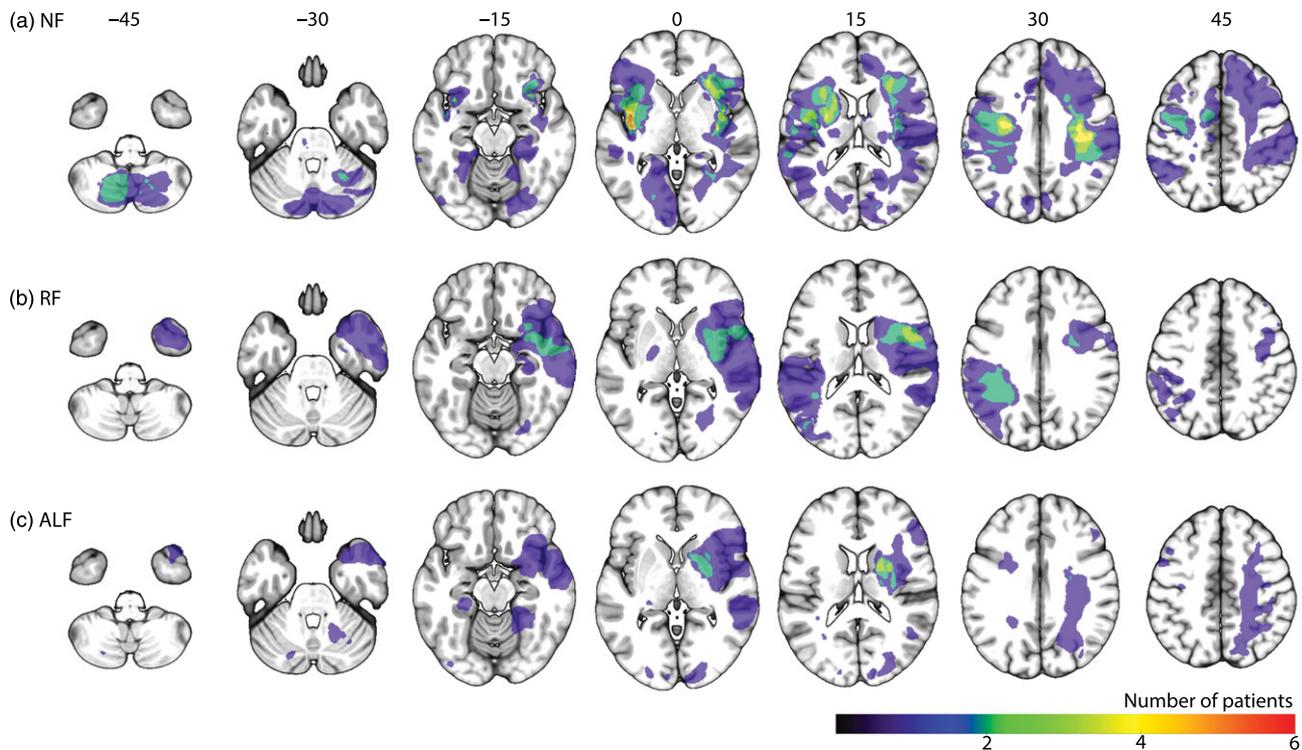
Figure 4. Mean confidence ratings (± standard deviations) for the controls, the no-forgetting patients (NF), the patients showing rapid forgetting (RF) and those showing accelerated long-term forgetting (ALF) patients immediately after the encoding phase (T1), after 20–30 min (T2) and after one week (T3). The slope of each line reflects an increase in confidence based on correct versus incorrect responding.

problem, most studies assessed ALF after optimizing initial learning (Geurts et al., 2015), also called "learning to criterion". However, our data show that initial learning does not substantially differ between our four groups. Therefore, we suggest that this potential confound has not affected the results of our study. Furthermore, a disadvantage of learning to criterion is that it may result in overlearning of the material, which, in turn, can lead to ceiling performances that mask early forgetting (Bell, 2006). In addition, optimizing initial learning does not reflect the learning and memory demands of most everyday situations (Geurts et al., 2019).

Future research should focus on functional connectivity to investigate networks underlying ALF and RF. In line with our suggestion that ALF and RF are two distinctive processes, relying on different parts of the brain, we expect disruptions in a distributed cortical network to be associated with ALF, while reduced functional connectivity between hippocampal areas and the cortex might be associated with RF. Based on the prevalence of both ALF and RF in stroke patients, stroke patients seem to be a suitable population for such network analysis.

To conclude, ALF often remains undetected in clinical practice in which the standard delay for testing is typically 20–30 min. This study complements previous research demonstrating that ALF patients show clearly distinguishable forgetting patterns compared to RF patients. ALF and RF seem to be two functionally distinct deficits. With this new set-up

of the Doors, wherein patients can perform the one-week delay in their home environment, ALF measurement becomes feasible for clinical practice.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617721000527

CONFLICTS OF INTEREST

The authors have no competing interests to disclose.

DATA AVAILABILITY STATEMENT

All data are stored on a University of Amsterdam server for the project FAB4V. Anonymized data are available upon request from the project leader (e.h.f.dehaan@uva.nl) one year after completion of the ERC project (01–07–2021).

REFERENCES

Adhikari, M. H., Hacker, C. D., Siegel, J. S., Griffa, A., Hagmann, P., Deco, G., & Corbetta, M. (2017). Decreased integration and information capacity in stroke measured by whole brain models of resting state activity. *Brain*, *140*, 1068–1085. https://doi.org/10.1093/brain/awx021

Baddeley, A., Atkinson, A., Kemp, S., & Allen, R. (2019). The problem of detecting long-term forgetting: Evidence from the Crimes Test and the Four Doors Test. *Cortex*, *110*, 69–79. https://doi.org/10.1016/j.cortex.2018.01.017

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (2006). Doors and People: A test of visual and verbal recall and recognition. *Bury St.* Edmunds, UK: Thames Valley Test Company.

Bell, B. D. (2006). WMS-III logical memory performance after a two-week delay in temporal lobe epilepsy and control groups. *Journal of Clinical and Experimental Neuropsychology*, *28*, 1435–1443. https://doi.org/10.1080/13803390500434367

Blake, R., Wroe, S., Breen, E., & McCarthy, R. A. (2000). Accelerated forgetting in patients with epilepsy: Evidence for an impairment in memory consolidation. *Brain*, *123*, 472–483. https://doi.org/10.1093/brain/123.3.472

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols*, *1*, 2277–2281. https://doi.org/10.1038/nprot.2006.390

Butler, C. R., Bhaduri, A., Acosta-Cabronero, J., Nestor, P. J., Kapur, N., Graham, K. S., . . . & Zeman, A. Z. (2009). Transient epileptic amnesia: Regional brain atrophy and its relationship to memory deficits. *Brain*, *132*, 357–368. https://doi.org/10.1093/brain/awn336

Butler, C., & Zeman, A. (2008). Recent insights into the impairment of memory in epilepsy: Transient epileptic amnesia, accelerated long-term forgetting and remote memory impairment. *Brain*, *131*, 2243–2263. https://doi.org/10.1093/brain/awn127

Cassel, A., & Kopelman, M. D. (2019). Have we forgotten about forgetting? A critical review of 'accelerated long-term forgetting' in temporal lobe epilepsy. *Cortex*, *110*, 141–149. https://doi.org/10.1016/j.cortex.2017.12.012.

Cassel, A., Morris, R., Koutroumanidis, M., & Kopelman, M. (2016). Forgetting in temporal lobe epilepsy: When does it become accelerated? *Cortex*, *78*, 70–84. https://doi.org/10.1016/j.cortex.2016.02.005

Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., & Friston, K. (2007). Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *Neuroimage*, *37*, 866–875. https://doi.org/10.1016/j.neuroimage.2007.04.065

De Graaf, A., & Deelman, B. G. (1991). *Cognitieve Screening Test* [Cognitive Screening Test]. Lisse, the Netherlands: Swets & Zeitlinger.

De Jonghe, J.F.M., Schmand, B., Ooms, M., & Ribbe, M. (1997). Verkorte Informantvragenlijst over cognitieve achteruitgang bij ouderen [Abbreviated form of the Informant Questionnaire on cognitive decline in the elderly]. *Tijdschrift voor Gerontologie en Geriatrie*, *28*, 224–229.

Dudai, Y. (2012). The restless engram: Consolidations never end. *Annual Review of Neuroscience*, *35*, 227–247. https://doi.org/10.1146/annurev-neuro-062111-150500

Duits, A., & Kessels, R. (2014). Schatten van premorbide functioneren [Estimating premorbid functioning]. In M. Hendriks, R. Kessels, M. Gorissen, B. Schmand, & A. Duits (Eds), *Neuropsychologische diagnostiek: De klinische praktijk* (pp. 173–186). Amsterdam, the Netherlands: Boom.

Elliott, G., Isaac, C. L., & Muhlert, N. (2014). Measuring forgetting: A critical review of accelerated long-term forgetting studies. *Cortex*, *54*, 16–32. https://doi.org/10.1016/j.cortex.2014.02.001

Fitzgerald, Z., Thayer, Z., Mohamed, A., & Miller, L. A. (2013). Examining factors related to accelerated long-term forgetting in epilepsy using ambulatory EEG monitoring. *Epilepsia*, *54*, 819–827. https://doi.org/10.1111/epi.12090

Geurts, S., van der Werf, S. P., & Kessels, R. P. C. (2015). Accelerated forgetting? An evaluation on the use of long-term forgetting rates in patients with memory problems. *Frontiers in Psychology*, *6*, 752. https://doi.org/10.3389/fpsyg.2015.00752

Geurts, S., van der Werf, S. P., Kwa, V. I., & Kessels, R. P. C. (2019). Accelerated long-term forgetting after TIA or minor stroke: A more sensitive measure for detecting subtle memory dysfunction? *Cortex*, *110,* 150–156. https://doi.org/10.1016/j.cortex.2018.04.002

Hochstenbach, J., Mulder, T., van Limbeek, J., Donders, R., & Schoonderwaldt, H. (1998). Cognitive decline following stroke: A comprehensive study of cognitive decline following stroke. *Journal of Clinical and Experimental Neuropsychology*, *20*, 503–517. https://doi.org/10.1076/jcen.20.4.503.1471

Hoefeijzers, S., Dewar, M., Della Sala, S., Zeman, A., & Butler, C. (2013). Accelerated long-term forgetting in transient epileptic amnesia: An acquisition or consolidation deficit? *Neuropsychologia*, *51*, 1549–1555. https://doi.org/10.1016/j.neuropsychologia.2013.04.017

Kantner, J., & Dobbins, I. G. (2019). Partitioning the sources of recognition confidence: The role of individual differences. *Psychonomic Bulletin & Review*, *26*, 1317–1324. https://doi.org/10.3758/s13423-019-01586-w

Kantner, J., & Lindsay, D. S. (2014). Cross-situational consistency in recognition memory response bias. *Psychonomic Bulletin & Review*, *21*, 1272–1280. https://doi.org/10.3758/s13423-014-0608-3

Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition*, *28*, 92–107. https://doi.org/10.3758/BF03211579

Lezak, M. D., Howieson, D., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.

Lim, C., & Alexander, M. P. (2009). Stroke and episodic memory disorders. *Neuropsychologia*, *47*, 3045–3058. https://doi.org/10.1016/j.neuropsychologia.2009.08.002

Loftus, G. R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 397. https://doi.org/10.1037/0278-7393.11.2.397

Mayes, A. R., Hunkin, N. M., Isaac, C., & Muhlert, N. (2019). Are there distinct forms of accelerated forgetting and, if so, why? *Cortex*, *110,* 115–126. https://doi.org/10.1016/j.cortex.2018.04.005

Migo, E. M., Mayes, A. R., & Montaldi, D. (2012). Measuring recollection and familiarity: Improving the remember/know procedure. *Consciousness and Cognition*, *21*, 1435–1455. https://doi.org/10.1016/j.concog.2012.04.014

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H. O. (2012). Age-specific CT and MRI templates for spatial normalization. *Neuroimage*, *61*(4), 957–965. https://doi.org/10.1016/j.neuroimage.2012.03.020

Saczynski, J. S., Sigurdsson, S., Jonsdottir, M. K., Eiriksdottir, G., Jonsson, P. V., Garcia, M. E., . . . Gudnason, V. (2009). Cerebral

infarcts and cognitive performance: Importance of location and number of infarcts. *Stroke*, *40*, 677–682. https://doi.org/10.1161/STROKEAHA.108.530212

Schmand, B., Lindeboom, J., & Van Harskamp, F. (1992). *Nederlandse Leestest voor Volwassenen* [Dutch Adult Reading Test]. Lisse, the Netherlands: Swets & Zeitlinger.

Schouten, E. A., Schiemanck, S. K., Brand, N., & Post, M. W. (2009). Long-term deficits in episodic memory after ischemic stroke: Evaluation and prediction of verbal and visual memory performance based on lesion characteristics. *Journal of Stroke and Cerebrovascular Diseases*, *18*(2), 128–138. https://doi.org/10.1016/j.jstrokecerebrovasdis.2008.09.017

Sekeres, M. J., Moscovitch, M., & Winocur, G. (2017). Mechanisms of memory consolidation and transformation. In N. Axmacher & B. Rasch (Eds.), *Cognitive neuroscience of memory consolidation* (pp. 17-44): Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-45066-7_2

Snaphaan, L., & de Leeuw, F.-E. (2007). Poststroke memory function in nondemented patients: A systematic review on frequency and neuroimaging correlates. *Stroke*, *38*, 198–203. https://doi.org/10.1161/01.STR.0000251842.34322.8f

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Tuladhar, A. M., Snaphaan, L., Shumskaya, E., Rijpkema, M., Fernández, G., Norris, D. G., & de Leeuw, F. E. (2013). Default mode network connectivity in stroke patients. *PloS One*, *8*, e66556. https://doi.org/10.1371/journal.pone.0066556

UNESCO (2011). *International Standard Classification of Education (ISCED 2011)*. Montreal, QC, Canada: UNESCO Institute for Statistics.

Van Der Elst, W. I. M., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2006). Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society*, *12*, 80–89. https://doi.org/10.1017/S1355617706060115

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale - Fourth Edition (WAIS-IV)*. San Antonio, TX: Pearson.

Wei, C., Zhang, F., Chen, L., Ma, X., Zhang, N., & Hao, J. (2016). Factors associated with post-stroke depression and fatigue: Lesion location and coping styles. *Journal of Neurology*, *263*, 269. https://doi.org/10.1007/s00415-015-7958-2

Weintraub, S., Wicklund, A. H., & Salmon, D. P. (2012). The neuropsychological profile of Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, *2*(4), a006171. https://doi.org/10.1101/cshperspect.a006171

Wheeler, M., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory*, *11*, 571–580. https://doi.org/10.1080/09658210244000414

Wilkinson, H., Holdstock, J. S., Baker, G., Herbert, A., Clague, F., & Downes, J. J. (2012). Long-term accelerated forgetting of verbal and non-verbal information in temporal lobe epilepsy. *Cortex*, *48*, 317–332. https://doi.org/10.1016/j.cortex.2011.01.002

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, *31*, 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x