Cognition 170 (2018) 254–262

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Extremely long-term memory and familiarity after 12 years

Christelle Larzabal[a,b,*], Eve Tramoni[c,d], Sophie Muratot[a,b], Simon J. Thorpe[a,b], Emmanuel J. Barbeau[a,b]

[a] Université de Toulouse, UPS, Centre de Recherche Cerveau et Cognition, France
[b] CNRS, CerCo, Toulouse, France
[c] INSERM U 751, Marseille, France
[d] Aix-Marseille Université, Faculté de Médecine, Marseille, France

ARTICLE INFO

Keywords:
Long-term memory
Familiarity
Explicit memory
Consciousness
Reactivation
Inactive memory

ABSTRACT

In 2006 Mitchell demonstrated that implicit memory was robust to decay. He showed that the ability to identify fragments of pictures seen 17 years before was significantly higher than for new stimuli. Is this true only for implicit memory? In this study, we tested whether explicit memory was still possible for drawings (n = 144) that had been presented once or three times, two seconds each time on average, approximately 12 years earlier. Surprisingly, our data reveal that our participants were able to recognize pictures above chance level. Preserved memory was mainly observed in the youngest subjects, for stimuli seen three times. Despite the fact that confidence judgments were low, reports suggest that recognition could be based on a strong sense of familiarity. These data extend Mitchell's findings and show that familiarity can also be robust to decay.

1. Introduction

Implicit memory does not involve explicit or conscious recollection and is inferred from behavioral performance (Graf & Schacter, 1985; Roediger, 2012; Schacter, 1987). Mitchell (2006) published a report demonstrating that implicit memory could be preserved on the very long-term and over a much longer time period than the several months that had been previously investigated (Kolers, 1976; Mitchell & Brown, 1988; Sloman, Hayman, Ohta, Law, & Tulving, 1988). In his experiment, participants were better at identifying fragments of black-and-white line drawings they had seen 17 years before than new drawings, even though some of them could not remember having participated in the original study. This study shows the longest duration of implicit memories to date.

In contrast to this, direct and intentional reference is made to prior learning events in explicit memory tests (Graf & Schacter, 1985; Schacter, 1987; Voss, Baym, & Paller, 2008). This holds true for recognition memory tasks where participants have to determine whether an item has already been experienced or not. It is generally assumed that recognition in such tasks relies on two processes: familiarity, which is the ability to identify that an item has previously been encountered, and recollection, which involves the retrieval of contextual details associated with the item (Mandler, 1980).

Numerous studies have shown that explicit memory can extend to a few decades if not for a lifetime. The first empirical evidence was provided in the late 19th century by Galton (1879) who could retrieve past events cued by a list of words, some of these events belonging to his childhood. This task was later developed (Crovitz & Sciffman, 1974) and used in several studies to measure participants' autobiographical memory across the lifespan (e.g. Fitzgerald & Lawrence, 1984; Rubin & Schulkind, 1997) including early childhood (Crovitz & Quina-Holland, 1976). These studies confirmed the ability to maintain very long-term memories although the accuracy of such episodic memories is difficult to verify.

Other tasks were developed in which remote semantic information was assessed. For example, participants were tested on their memory of old classmates (Bahrick, Bahrick, & Wittlinger, 1975), knowledge learned in school (Conway, Cohen, & Stanhope, 1991) or old TV program titles (Squire, 1989). In all cases, participants were able to retrieve a significant amount of information.

These studies demonstrated that memories could survive for several decades despite performance decreasing over time. Indeed, performance quickly decreases after the acquisition phase before stabilizing several years later. This "forgetting curve" was first described by Ebbinghaus (1885) and seems to correspond to the power function of time according to recent models (Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991). Interestingly, the manipulation of different experimental parameters can affect the shape of the forgetting curve and slow down decay over time. Spaced repetition of new information leads to longer lasting memory traces than

* Corresponding author at: Université de Toulouse, UPS, Centre de Recherche Cerveau et Cognition, France.
E-mail address: christelle.larzabal@cnrs.fr (C. Larzabal).

http://dx.doi.org/10.1016/j.cognition.2017.10.009
Received 29 May 2017; Received in revised form 13 October 2017; Accepted 17 October 2017
0010-0277/

single exposure or massed repetition (Dempster, 1988; Ebbinghaus, 1885; Hintzman, 1976; Melton, 1970). Active reactivation of the stimuli through testing results in better performance than passive exposure (Carrier & Pashler, 1992; Roediger & Karpicke, 2006). Emotional content can also directly influence involuntary reactivations of an event and therefore its long-term maintenance (Berntsen & Rubin, 2002; Waters & Leeper, 1936).

Importantly, such reactivations slow down forgetting by directly influencing the neurobiological mechanisms involved in the consolidation of a new memory trace. Indeed, after the stabilization of its synaptic weights in localized networks, the memory trace requires consolidation at the system level (Frankland & Bontempi, 2005). During this phase, repeated reactivations of the memory trace allow the strengthening of the engram in the neocortical structures and its long-term maintenance (Girardeau, Benchenane, Wiener, Buzsáki, & Zugaro, 2009; Sutherland & McNaughton, 2000). Therefore, reactivation of the memory trace through direct re-exposure to the stimulus or when the stimulus is mentally evoked would help the completion of such a process.

From the literature reviewed above, it appears that information acquired from everyday experiences can last for a decade or even longer. In all these studies, the information was presented for at least several hours and repeated many times (TV programs, classmates' names and faces, conceptual knowledge, etc.). In addition, it was probably reactivated either voluntarily (knowledge learned in school) or involuntarily, for example if associated with strong emotions. All these factors increase the probability of retrieving information and therefore increase long-term maintenance. Hence, these conditions are very different from Mitchell's experiment where very long-term priming was found for simple drawings that had been presented for one to three seconds and at most three times. Therefore, we do not know how long explicit memory can be maintained if exposure to the material is brief and especially if no rehearsal has taken place in the meantime.

Recent neuronal models based on Spike Time Dependent Plasticity (Masquelier & Thorpe, 2007) show that a few exposures to simple stimuli can be enough to create hyper-selective neurons. Such neurons could potentially remain silent and their synaptic weights could be preserved until the same stimulus or a close representation is presented again (Barron et al., 2016). According to this view, implicit but also explicit memories might be maintained for a life-time (Thorpe, 2011). Although theoretical and based on modeling, this point of view provides plausible mechanisms suggesting that very long-term memory is possible in the absence of rehearsal. In this study we decided to test it experimentally.

In the present paper, we tested participants' recognition memory for simple colored drawings that had been presented once or three times for two seconds each time on average, between 8 and 14 years earlier. Because the stimuli were simple and not particularly interesting (a third of them were actually abstract figures which were difficult to verbalize), we thought that it was unlikely that the participants ever thought about them in the meantime.

2. Materials and methods

2.1. Participants

Between 2002 and 2008 (mean year = 2004, SD = 2.2), 243 healthy control subjects participated in a neuropsychological protocol in the Timone University Hospital in Marseille (France) with the aim of building normative data. This protocol included four cognitive tasks among which there was a visual recognition memory test that we used as the basis of the current study.

In 2016 we were able to retrieve telephone numbers and/or postal addresses for 63 of the initial 243 participants. These 63 subjects were called back or written to and informed that we would like them to come back for a scientific experiment to test their memory of the pictures presented in the initial test. They were also informed that we thought their memory could be maintained for such a long time and that this was precisely why we wanted to do the test. Importantly, we were not able to obtain written consent for 38 of these 63 individuals: 31 participants could not be contacted either because the contact details taken at the time of the initial test had changed or because they simply did not respond to our messages; two participants passed away during the intervening period and their widows, who also participated in the initial study, did not want to do the test; one participant moved away and could not come back to do the experiment; two participants overtly declined the invitation to participate in our experiment because of a lack of interest. Overall, twenty-five subjects agreed to participate. One of them was not included because of macular degeneration. Thus 24 subjects were included in this study. None reported the presence of important neurological, psychological or psychiatric troubles since 2002 and they are referred to as test participants in this paper (14 females, mean age = 60.8 y, SD = 17.5, range: 34–84 y). All test subjects had normal or corrected-to-normal vision and audition. The overall cognitive abilities of the test participants beyond age 60 were assessed in 2002–2008 using the Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) from the French GRECO consensual version (mean MMSE score 29.5 out of 30, SD = 0.7, range: 28–30). All test participants were tested with the MMSE again in the context of the current study with no deficit found for any of them (mean score = 29.5 out of 30, SD = 0.9, range: 27–30).

A group of 24 naïve control subjects, who had never seen any of the drawings from the first experiment and matched for gender and age, were also recruited (14 females, mean age = 62.5 y, SD = 20.7, range: 33–94 y).

The current study was approved by the INSERM Ethical Evaluation Committee (CEEI, N°15-263) and all participants gave informed written consent before experimentation.

2.2. Initial test: the DMS-48

The basis of the current study is a visual recognition memory test, the DMS-48, that is widely used in memory clinics in French-speaking countries (see for example Barbeau et al., 2004; Didic et al., 2013). The DMS-48 is freely available for research purposes (http://cerco.ups-tlse.fr/~barbeau/dms48.html).

During an incidental encoding phase (test participants were not told that it was a memory test), 48 pictures were presented one after the other. For each stimulus, test participants were simply instructed to say if the picture had more or less than three colors. Two forced-choice recognition memory phases took place after three and 60 min. Different neuropsychological tasks were presented during the interfering phases. Test participants were asked to identify which picture they had previously seen by choosing from two stimuli presented simultaneously (one old stimulus and one distractor each time). There were three categories (examples in Fig. 1). In the Unique category, the old and distractor pictures were real-world objects with names and shapes (e.g. a rooster and a digger). In the Paired category, the old and distractor items had the same name and similar shapes and colors (e.g. two different snowmen). The Abstract category involved abstract patterns that were difficult to verbalize. A separate set of 48 new distractors was used for each of the test phases. All stimuli were presented on A4 white sheets of papers.

Performance on the DMS-48 is expressed as a percentage (maximum performance 100%).

2.3. Initial number of presentations of the stimuli

Overall, the 48 old stimuli were seen three times (during the incidental encoding phase and the recognition phases at three and 60 min), whereas the other 96 stimuli that served as distractors during the recognition phases at three and 60 min were seen only once. This

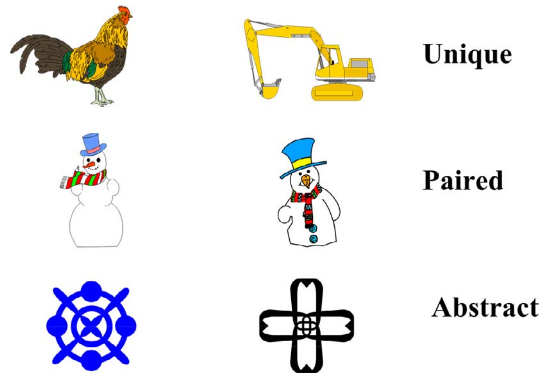


Fig. 1. Examples of pairs of stimuli used in the DMS-48 for the three categories.

was the case for 19 test subjects because the remaining five test subjects participated in other experimental designs involving longer retention intervals of the material. For this reason, two subjects (participant 4 aged 64 and participant 5 aged 67) underwent a third recognition phase after six weeks in 2007 and were thus exposed to the old stimuli four times. The remaining three subjects (participant 2 aged 83, participant 3 aged 84 and participant 6 aged 63) underwent the whole protocol again 18 months later and were thus exposed to the old stimuli six times and to the distractors twice (see Supplementary information, Table S1).

2.4. Initial time spent on the stimuli

The time spent on performing the incidental encoding of the 48 stimuli and completing each of the two recognition tests was calculated for 19 out of the 24 participants. For the remaining five participants the time spent on performing the two recognition tests was not measured.

On average participants took 212 s (SD = 42.5, range: 150–327 s) to perform the incidental learning task, 166 s (SD = 52, range: 83–310 s) to complete the first recognition test and 159 s (SD = 47.4, range: 76–279 s) to finish the second recognition test. By dividing the duration of these three experimental tasks by the number of stimuli presented, we were able to get an estimate of the exposure time for each of the 144 pictures. On average, the 48 stimuli presented three times were seen for 7.8 s (SD = 1.7, range: 5–12.9 s), whereas the 96 stimuli presented once were seen for 1.7 s (SD = 0.5, range: 0.8–3.2 s). Note that this calculation has some limitations, given that decisions on a forced-choice task can be made after considering only one of the two items (Starns, Chen, & Staub, 2017).

2.5. Initial results

The test participants' performance reached a ceiling (after three minutes: mean = 98.4%, SD = 2.0, range: 92–100% and after 60 min, mean = 98.5%, SD = 1.9, range: 92–100%) thus, demonstrating that essentially all the targets had been fully processed and memorized for at least an hour.

2.6. Current tasks

In the present study, we used two successive recognition memory tasks to assess the test participants' ability to recognize the stimuli seen in the initial experiment. The first one was a forced-choice task which was as close as possible to the initial one, but using a new set of distractor stimuli. The second one was a Yes/No task to assess whether the findings could be replicated in a different context (Fig. 2).

As far as we know this is the first study of its kind. As such we tried to ensure that environmental variables were the same for the initial exposure and the present experiment, thus trying to respect the encoding specificity principle (Tulving & Thomson, 1973). We used a similar method to present the stimuli (printed on A4 white sheets presented in binders) and subjects were seen in the same hospital.

The forced-choice task was based on 144 pairs of colored drawings. For the three stimulus categories: Unique, Paired, Abstract, each pair was made of an old stimulus seen in the initial task and a new one. Old stimuli were either the 48 targets seen several times or the 96 distractors seen only once in the initial test. The new stimuli were chosen with respect to the original combination of categories Unique, Paired, and Abstract (n = 48 for each). Old and new items were shuffled within the Unique and Abstract categories, resulting in different pairs of Unique and Abstract stimuli across participants. For the Paired category, the same old/new combination was presented for all participants. For each trial, the two items were displayed on the right and left-hand side of the sheet (counterbalanced). The order of the presentation of the pairs of the stimuli was randomized across participants. The forced-choice task was divided into three blocks of 48 pairs of pictures. Within each block, trials were pseudo-randomized to equalize the number of stimuli seen three times and once as well as the number of stimuli belonging to each combination category (Unique, Paired and Abstract).

For each trial, test participants were invited to identify the picture that was presented in the initial experiment several years before.

This task was also presented to the 24 naive control subjects. Of course, they had not seen any of the stimuli before, so they were asked to say which one they thought had been presented previously. This procedure was performed because the new stimuli used in the forced-choice were selected in 2016. There could thus have been a risk that these new items were inadvertently systematically different from the previous ones, which would have introduced a bias in the experiment.

After completing the forced-choice task, the test participants were invited to perform a Yes/No recognition memory. In this task, 144 out of the 288 stimuli previously shown were individually presented, centered on A4-printed white sheets. These stimuli were equalized across the three categories Unique, Paired and Abstract. They included the stimuli seen three times (n = 48) and once (n = 48) in the initial test as well as the new stimuli used in the previous forced-choice task (n = 48). Only 48 stimuli were used for each category to limit the duration of the Yes/No task. Thus, all stimuli used in the Yes/No task had been presented in the forced-choice task.

For each stimulus, test participants were asked to respond "yes" if the picture was presented in the initial experiment or "no" if it was a new picture which they had just seen in the previous forced-choice task.

Two participants refused to perform this task. One because she said she saw too many similar pictures in her life (she used to work with children in a day nursery) and the other one because it was too hard for him to differentiate between the drawings seen in the initial test and those which he had just seen.

2.7. Confidence responses and verbal reports

For each trial in the forced-choice and the Yes/No tasks test participants were asked to rate their confidence in their response on a 5-point scale. Here, the original French terms used are presented in italics along with their English translation in brackets; 0: '*au hasard*' ('guess'), 1: '*pas du tout sûr*' ('not sure at all'); 2: '*peu sûr*' ('not very sure'); 3: '*plutôt sûr*' ('fairly sure'); 4: '*sûr*' ('sure'). Importantly, if the test participants used a confidence score of 3 or 4, they were given the opportunity to explain their choice. Indeed, given that high confidence responses can be based on recollection, we wanted to assess whether test participants would spontaneously recollect elements of the initial test that would help them recognize some stimuli.

2.8. Performance and presentation of the results

Performance in the forced-choice task was expressed as a percentage (maximum performance 100%).

In the Yes/No task, performance was calculated using a sensitivity index: d-prime = Z(Hit rate) − Z(False Alarm rate) with Z being a Z-

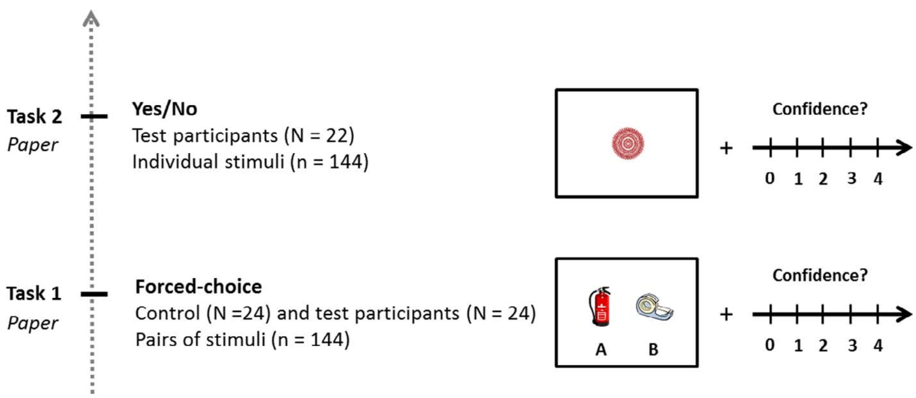


Fig. 2. The current experimental procedure consisted of two explicit memory tests: a forced-choice and a Yes/No recognition task. In the forced-choice task, participants were asked to identify which of the two stimuli was presented in the initial experiment. In the Yes/No task test participants were asked to decide whether the individual stimuli corresponded to 'old' or 'new' items. In both tasks each response was followed by a confidence score rated on a 5-point scale.

score. Within each analysis, Z-score values were corrected for infinite scores (in accordance with Snodgrass & Corwin, 1988). A d-prime value of 0 corresponds to chance level performance and larger d-prime values correspond to above chance level performance.

To get a better view of the data, the estimated number of participants for a given correct percentage was also calculated using a kernel density estimation (Allen, Erhardt, & Calhoun, 2012). This allowed us to get a continuous representation of the data in the forced-choice and Yes/No tasks that cannot be visualized via classical bar or box plots.

2.9. Statistical analyses

Chance level on the forced-choice was calculated based on the estimate of a binomial distribution (50%) which varies according to the number of trials. The maximum chance corresponded to the upper 95% confidence interval of the binomial distribution. Therefore, for the 144 pairs of drawings presented in the forced-choice task, participants would perform better than the chance level if they were more than 58.4% correct. This means that the participants would correctly identify more than 84 items (144/2 + 95%CI). Repeating such a measure across participants could lead to some uncertainty of this estimate at the group level. To be at chance level, we considered that 95% of a group of participants should fall below the maximum chance. For 24 participants, this means that 22–23 participants should correctly identify at most 84 pictures and only one or two participants could recognize more items. In the case where more than two participants have performance beyond the maximum chance, the population deviates from chance level.

3. Results

Only the results of the forced-choice recognition (first task) are presented in this section. The purpose of the Yes/No recognition (second task) was to test whether the results in the forced-choice could be replicated in a different context. This is not of primary relevance. Moreover, the responses in the Yes/No task might have been directly influenced by the judgments made at first on the forced-choice task. For these reasons, the results of the Yes/No task are not described in the main manuscript but can be found in the Supplementary information section.

3.1. Naïve control participants' performance

The performance of naive control subjects who had never seen any of the stimuli was compared to the average and maximum chance level. Control participants were on average 50.9% correct in the forced-choice task (mean, SD = 5.9, range 41.7–63.2%) with only two participants performing above the maximum chance limit.

3.2. Test participants' performance

In the forced-choice task, two drawings were simultaneously presented: one corresponding to a picture presented in the initial experiment and one corresponding to a new picture. Interestingly, the median distribution of test participants' performance was significantly different from the control group performance (55.2%, vs 51.4%, $U = 387$, $CI_{95} = [0–8.3]$, Mann-Whitney-Wilcoxon-test). This shows that the whole population of test participants deviated from chance level (Fig. 3). Test participants' mean performance on the forced-choice was 55.3% (SD = 7.2, range: 40.3–69.4%) with eight participants above the maximum chance limit of 58.4%. This means that the eight best participants could identify between one and 15 pictures (mean = 7.4, SD = 3.9) more than the maximum 84 pictures expected by chance.

3.3. Effect of the number of exposures

In the initial test, 48 stimuli were seen three times (= old stimuli at the time) and 96 only once (= distractors at the time). This was true for all of the participants except for two (participants 4 and 5) who were exposed four times to the old stimuli and three others (participants 2, 3 and 6) who were exposed six times to the old stimuli and two times to the distractors. In the forced-choice task, two of the subjects who had

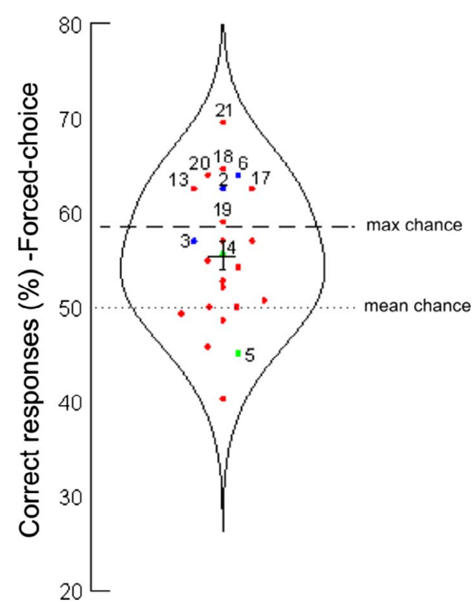


Fig. 3. Test participants' performance on the forced-choice recognition task (crosshair: mean; surrounding shape: continuous estimate). One dot is the performance of one participant. The participant numbers for each of the eight best participants as well as those who saw the old stimuli four times (green dots) or six times (blue dots) are shown on the figure.

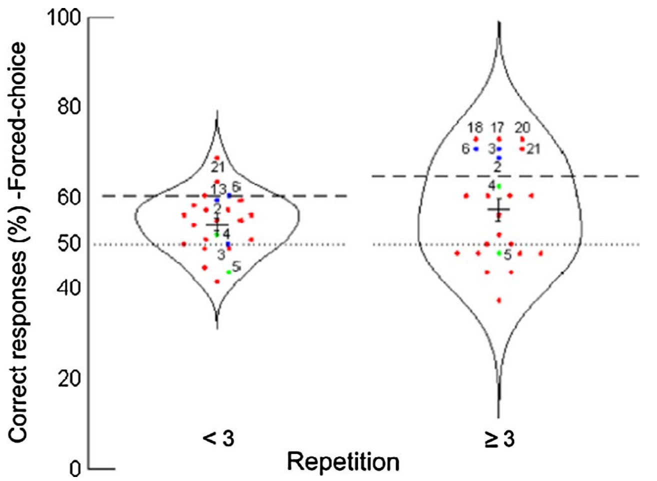


Fig. 4. Test participants' performance according to the number of times the stimuli were seen. See caption in Fig. 3 for more details.

seen the stimuli more times than the others were among the eight best subjects mentioned above (participants 2 and 6). Overall, the 24 participants were on average 57.4% correct (SD = 11.0, range: 37.5–72.9%) for stimuli seen at least three times and 54.3% correct (SD = 6.4, range: 41.7–68.7%) for stimuli seen fewer than three times (Fig. 4). Seven test participants were above the maximum chance limit of 64.8% for stimuli seen at least three times. These participants were able to correctly identify between two and four pictures more than the maximum 31 pictures expected by chance level. Four of them corresponded to participants who saw the stimuli only three times and three of them to the participants who saw the stimuli six times. In comparison, only two test participants were above the maximum chance limit of 60.4% for stimuli seen fewer than three times.

3.4. Stimulus type

This experiment used stimuli from three categories: Unique, Paired and Abstract. Test participants' performance did not differ over the three categories.

3.5. Other influencing factors

We tested the potential influence of several general factors in the forced-choice task. We found that participants' performance was negatively correlated with their age ($r = -0.44$, $CI_{95} = [-0.7 \text{ to } 0.0]$, Pearson's correlation coefficient), but not with gender ($r = -0.11$, $CI_{95} = [-0.49 \text{ to } 0.3]$, Pearson's correlation coefficient) or educational level ($r = -0.0$, Pearson's correlation coefficient).

3.6. Confidence rating

Every response was followed by a 5-point scale confidence level ranging from 0: 'guess' to 4: 'sure'. On average test participants were not sure of their responses (mean = 1.0, SD = 0.9, range: 0–2.4) with 46.8% (SD = 41.1) of the responses rated as a 'guess' (rate 0) and 23.3% (SD = 31.1) as 'not sure at all' (rate 1).

Performance increased with confidence (Fig. 5) but this increase was not significant (One-way ANOVA: $F(4,71) = 1.12$; CI_{95}-difference (Guess vs Not sure at all) = [−20.6 to 19.6], CI_{95}-difference (Guess vs Not very sure) = [−30.6 to 10.8], CI_{95}-difference (Guess vs Fairly sure) = [−38.3 to 11.6], CI_{95}-difference (Guess vs Sure) = [−27.9 to 37.1]).

The mean confidence of the eight best participants was low (mean = 0.7, SD = 0.9) except for participant 2 who reported the highest score out of all the test participants: mean = 2.4.

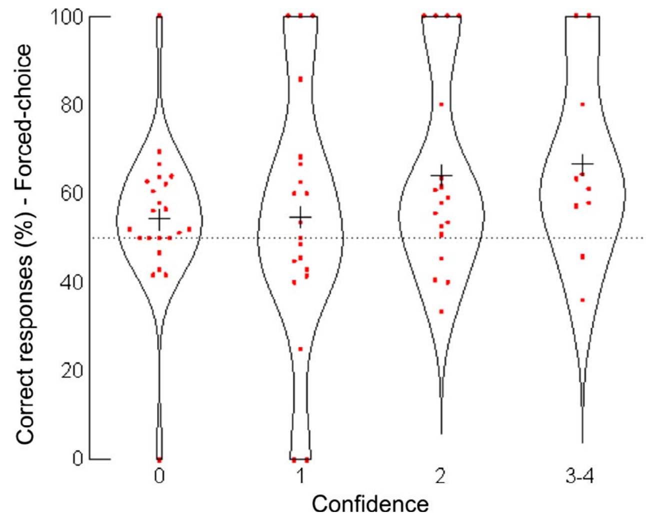


Fig. 5. Test participants' performance over the confidence level. 'Fairly sure' and 'sure' responses were grouped together because there were few occurrences. See Fig. 3 caption for more information.

3.7. Test participants' reports

Test participants were invited to provide a justification when they were 'fairly sure' (rate 3) or 'sure' (rate 4) of their responses. To ensure that there were no artificial justifications, they were not obliged to provide a response. The proportion of responses rated as 'fairly sure' and 'sure' was respectively 9.7% (SD = 16.8) and 1.1% (SD = 3.2). Overall, 61.7% (SD = 32.1, range: 19.7–100%) of the responses with a confidence level of 3 or 4 were associated with some verbal reports. Drawings associated with comments were correctly identified 71.4% (SD = 17.9, range: 46–100%) of the time. Importantly, these justifications referred to visual properties of the stimuli or familiarity with the items. However, none of the reports referred to contextual details which would have been indicative of recollection (see Table 1 for representative examples from three participants who saw the stimuli three times).

3.8. Test participants' recollection of the initial test

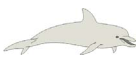
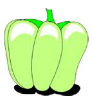
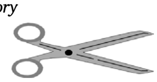
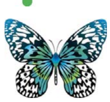
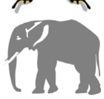
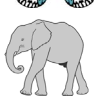
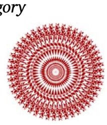
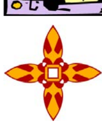
Two test participants out of 22 who were asked whether they recollected the initial test did not remember that they had performed it several years before. Interestingly, one of them (participant 18) was one of the eight best performers. None of the 22 test participants reported seeing or rehearsing any of the stimuli presented in the intervening time.

4. Discussion

In this study we found that our group of test participants was able to recognize simple colored pictures seen for a few seconds between eight and 14 years earlier. Our best performer, who had been exposed to the pictures at most three times, was able to identify 15 pictures more than the 84 pictures expected by chance. Note that no instruction to learn the stimuli was ever given to the subjects, even at initial encoding, which makes this performance even more remarkable. This extends Mitchell's findings and shows that explicit memories can be maintained for a decade.

Familiarity is thought to be a process used in long-term memory but it has never been demonstrated that it could be used for very long retention intervals. Our data provide strong evidence that it is the case. Indeed, an analysis of the verbal reports provided for high confidence responses revealed that details (shape or color of the stimuli) or feelings about the pictures ("it rings a bell", or participants "thought" they had

Table 1
Examples of reports for stimuli seen three times.

Participant	Stimulus condition		Confidence	Participant report	
	Three times	New		English translation	Original French
Unique category					
1			3	It rings a bell	Ça me parle
1			3	The drawing looks more familiar than the other one	Le dessin me parle plus que l'autre
11			4	I cannot remember seeing the sweater at all	Je ne me souviens pas du tout d'avoir vu le sweat-shirt
11			4	I remember seeing marine animals but not the dolphin specifically	Je me souviens d'avoir vu des animaux marins, mais pas spécifiquement du dauphin
11			4	By process of elimination	Par elimination
11			3	The shape of the building makes me think that I saw this drawing	La forme du bâtiment me fait penser que j'ai vu cette image
16			3	I think I have seen it	Il me semble l'avoir vu
Paired category					
1			3	I do not know the reason why	Je ne peux pas dire la raison
1			3	I do not think I saw anything symmetrical	Je ne pense pas avoir vu quelque chose de symétrique
1			3	Because of the color	Question de couleur
11			3	Because of the originality of the drawing	A cause de l'originalité du dessin
11			3	The shape seems more real	Car la forme semble plus réelle
Abstract category					
1			3	Because of the color	A cause de la couleur
11			3	By default because I do not remember the other picture	Par défaut car je ne me souviens pas de l'autre image
16			3	It looks quite familiar	Ça me parle bien

seen the picture) helped recognition. These responses were never associated with the recollection of contextual details or any reference to the initial encoding stage, which is typical of Remember responses. Instead, they correspond to Know judgments according to the dual-process theory (Mandler, 1980). In recognition memory tasks, Know judgments are usually associated with low confidence levels and Remember judgments with high confidence levels (Yonelinas, 2001) but as we found in our data this is not always the case. Indeed, in this study participants relied on familiarity even for the highest levels of confidence (levels 3 and 4). This could simply reflect the continuous nature of both recollection and familiarity (Wixted, 2009). Interestingly, some subjects also appeared to assess the novelty of the distractor in the forced-choice task, suggesting that correct recognition might have relied on a balance between familiarity and novelty assessment in some instances (Delhaye, Bastin, Moulin, Besson, & Barbeau, 2017). Most of the items (∼85%) however were recognized with low confidence levels suggesting that participants were unsure of their decisions. Although automatic processes such as perceptual fluency might influence retrieval (Jacoby, 1991), many of these responses can be viewed as low-confidence Know judgments relying on familiarity (Besson et al., 2015; Jacoby & Whitehouse, 1989; Voss et al., 2008). This was clearly the case for a few confidence level 2 responses ('not very sure'). Indeed, participants spontaneously reported familiarity even though they were not asked to give reasons for their choices (e.g. participant 4: "The indented side of the outer part reminds me of something" ("*le coté dentelé à la périphérie m'évoque quelque chose*"); participant 7: "feeling of familiarity" ("*sentiment de familiarité*"); participant 16: "It rings a bell" ("*Ça me parle*").

Importantly, a third of the test participants (eight out of 24) performed above chance level. However, these eight best performers were identified using a strict criterion. Therefore, we cannot completely rule out the possibility that other participants were able to retrieve more items than chance level. Indeed, the whole population of test participants clearly deviated from chance level and the distribution of the performance did not reveal obvious subgroups (i.e. those who could recognize the stimuli and those who could not). Several factors can explain the results found at the group level. In the following sections we discuss the effect of age, environmental context, subject selection and stimulus type.

We found that performance was negatively correlated with participants' age. This is not surprising given that aging is associated with a decline in the ability to encode and explicitly retrieve new stimuli (Light, 1991; Mitchell, 1989). Therefore, participants who were in their twenties during the initial test are expected to perform better than participants who were in their sixties. This is consistent with our data where the five youngest participants, who were between 20 and 23 years old at the time of acquisition of the stimuli, were in the top eight best performers. However, given the high correlation between the age at encoding and the age at retrieval it is difficult to know which of the two had more of an effect on participants' performance.

As opposed to priming, it has been shown that retrieval during explicit memory tests is influenced by the environmental context (McKone & French, 2001). We were careful in trying to keep these environmental variables the same from the initial exposure to the present experiment, thus respecting the encoding specificity principle (Tulving & Thomson, 1973): subjects were tested in the same hospital and stimuli were presented on similar white sheets in similar binders using a similar forced-choice task. Observing this encoding specificity might have helped participants to retrieve the items. However, participants' performance was still above chance level in the second task they performed: the Yes/No task, which was different to the initial test (see Supplementary information). Therefore, it seems likely that memory can be robust to experimental changes. Nevertheless, given that the two tasks were done successively some biases might have been included. For example, the judgments made at first on the forced-choice task might have influenced the responses in the Yes/No task. This could explain the correlation between participants' performance in the forced-choice and the Yes/No task. Further designs should avoid such potential biases to ensure robust conclusions.

Importantly, 39 out of the 63 individuals for whom we had contact details were not able to perform this study. This was mainly due to unfortunate reasons rather than overt refusal. It is important to note that even the 24 participants who agreed to participate were very skeptical concerning their memory abilities. They participated in the study out of good will with no expectation of performing well in the tests. Two of them were not able to remember that they performed the initial test but in spite of this, one of them was among the eight best performers. Given this, we do not know to what extent the 24 participants that we tested had overall better memory performance than the participants who did not perform the study.

In this study we also tested the recognition of simple colored-drawings. Evidence has shown that memory of pictures is better than words (Paivio & Csapo, 1973), a phenomenon known as the "picture superiority effect". This effect relies on the notion of distinctiveness where pictures can be differentiated from each other given their inherent features (Mintzer & Snodgrass, 1999). Each item would therefore be unique and could be differentiated from highly similar pictures as it has been shown (Brady, Konkle, Gill, Oliva, & Alvarez, 2013). High distinctiveness would prevent potential interferences and would explain why pictorial memories can be explicitly retrieved after a few days (Standing, 1973), a year (Nickerson, 1968) or even after a decade in our experiment. They might constitute "episodic elements" that would help form an episodic memory if combined with contextual details (Conway, 2009). Surprisingly, our data do not show any clear difference between the three stimulus categories: Unique, Paired and Abstract. This reinforces the idea that what we tested is mainly a pictorial memory rather than a combination of verbal or semantic and visual memory. Therefore, our results might only hold true for pictures.

In the initial test, participants reached a ceiling at 3 min and an hour after the encoding. Other paradigms using the DMS-48 also assessed participants performance within a week (Vallet et al., 2016) or six weeks of retention interval (data not published). Interestingly, performance was still on average above 90% for these longer periods. This small decline in the first two months following encoding fits well with previous studies that have investigated the time-course of the forgetting rate for picture recognition (Gehring, Toglia, & Kimble, 1976; Lawless, 1978). According to these data, recognition performance could be modeled using the power law function (see Rubin & Wenzel, 1996 for the best fit function calculation using Gehring et al., 1976 dataset). The shape of such a function would then explain why participants are about 60% correct after a year of retention interval (Nickerson, 1968; Shepard, 1967) and why 12 years later, participants are still above chance level in our study.

A crucial question that remains to be answered is how is it biologically possible that simple pictures that have been presented just three times and for only a few seconds, are still able to drive explicit choices a decade later? These memories might have been left in a dormant state (Dudai, 2004; Lewis, 1979) until they were woken-up through the presentation of the stimuli. Certain neuronal models using Spike Time Dependent Plasticity (STDP) mechanisms have shown that a few exposures are enough to create hyper-selective neurons (Masquelier & Thorpe, 2007). It remains to be clarified how these models relate to our experimental data.

5. Conclusions

We show here that three presentations of a stimulus might be enough to create a familiar representation in very long-term memory, particularly in young subjects. The precise mechanisms allowing such a feat remain to be clarified.

Author contributions

Conceptualization and methodology: C.L., E.J.B., E.T., S.J.T.; Investigation: E.T. and S.M.; Writing: C.L., E.J.B., S.J.T; Funding acquisition: S.J.T; Supervision: E.J.B.

Acknowledgments

This research was supported by the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n°323711 to S.J.T. The authors also acknowledge the support of the

French Agence Nationale de la Recherche (ANR) under grant ANR-12-BSH2-0010 (project ELMA). We would like to thank Jean-Michel Hupé for his constructive feedback in the statistical analysis of the data and Djamaa Atamena for her help in the preparation of the stimuli.

Appendix A. Supplementary material

The material and the data used are freely available in a Mendeley Data archive (Larzabal, Tramoni, Muratot, Thorpe, & Barbeau, 2017) located at http://doi.org/10.17632/dzn5wxyxgz.1. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cognition.2017.10.009.

References

Allen, E. A., Erhardt, E. B., & Calhoun, V. D. (2012). Data visualization in the neurosciences: Overcoming the curse of dimensionality. *Neuron, 74*(4), 603–608. http://dx.doi.org/10.1016/j.neuron.2012.05.001.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*(6), 396–408.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54–75. http://dx.doi.org/10.1037//0096-3445.104.1.54.

Barbeau, E., Didic, M., Tramoni, E., Felician, O., Joubert, S., Sontheimer, A., ... Poncet, M. (2004). Evaluation of visual recognition memory in MCI patients. *Neurology, 62*(8), 1317–1322. http://dx.doi.org/10.1212/01.WNL.0000120548.24298.DB.

Barron, H. C., Vogels, T. P., Emir, U. E., Makin, T. R., O'Shea, J., Clare, S., ... Behrens, T. E. (2016). Unmasking latent inhibitory connections in human cortex to reveal dormant cortical memories. *Neuron, 90*(1), 191–203. http://dx.doi.org/10.1016/j.neuron.2016.02.031.

Berntsen, D., & Rubin, D. C. (2002). Emotionally charged autobiographical memories across the life span: The recall of happy, sad, traumatic and involuntary memories. *Psychology and Aging, 17*(4), 636–652. http://dx.doi.org/10.1037//0882-7974.17.4.636.

Besson, G., Ceccaldi, M., Tramoni, E., Felician, O., Didic, M., & Barbeau, E. J. (2015). Fast, but not slow, familiarity is preserved in patients with amnestic mild cognitive impairment. *Cortex, 65*(November 2014), 36–49. http://dx.doi.org/10.1016/j.cortex.2014.10.020.

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual long-term memory has the same limit on fidelity as visual working memory. *Psychological Science, 24*(6), 981–990. http://dx.doi.org/10.1177/0956797612465439.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20*(6), 633–642. http://dx.doi.org/10.3758/BF03202713.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia, 47*(11), 2305–2313. http://dx.doi.org/10.1016/j.neuropsychologia.2009.02.003.

Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General, 120*(4), 395–409.

Crovitz, H. F., & Quina-Holland, K. (1976). Proportion of episodic memories from early childhood by years of age. *Bulletin of the Psychonomic Society, 7*(1), 61–62. http://dx.doi.org/10.3758/BF03337122.

Crovitz, H. F., & Sciffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society, 4*(5B), 517–518. http://dx.doi.org/10.3758/BF03334277.

Delhaye, E., Bastin, C., Moulin, C. J., Besson, G., & Barbeau, E. J. (2017). Bridging novelty and familiarity-based recognition memory as a single process: A matter of timing. *Visual Cognition,* 1–7.

Dempster, F. N. (1988). The spacing effect. *American Psychologist, 43*(8), 627–634.

Didic, M., Felician, O., Barbeau, E. J., Mancini, J., Latger-Florence, C., Tramoni, E., & Ceccaldi, M. (2013). Impaired visual recognition memory predicts Alzheimer's disease in amnestic mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders, 35*(5–6), 291–299. http://dx.doi.org/10.1159/000347203.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annual Review of Psychology, 55*, 51–86. http://dx.doi.org/10.1146/annurev.psych.55.090902.142050.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology.* New York, NY, US: Teachers College Press.

Fitzgerald, J. M., & Lawrence, R. (1984). Autobiographical memory across the adult lifespan. *Journal of Gerontology, 39*(6), 692–698.

Folstein, M., Folstein, S., & McHugh, P. R. (1975). A practical state method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience, 6*(2), 119–130. http://dx.doi.org/10.1038/nrn1607.

Galton, F. (1879). Psychometric experiments. *Brain, 2*, 148–162.

Gehring, R. E., Toglia, M. P., & Kimble, G.a. (1976). Recognition memory for words and pictures at short and long retention intervals. *Memory & Cognition, 4*(3), 256–260. http://dx.doi.org/10.3758/BF03213172.

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., & Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience, 12*(10), 1222–1223. http://dx.doi.org/10.1038/nn.2384.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 11*(3), 501–518. http://dx.doi.org/10.1037/0278-7393.11.3.501.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Vol. Ed.), *The psychology of learning and motivation: Vol. 7421*, (pp. 47–91). New York: Academic Press http://dx.doi.org/10.1016/S0079-7421(08)60464-8.

Jacoby, L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Jacoby, L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General, 118*(2), 126–135. http://dx.doi.org/10.1037//0096-3445.118.2.126.

Kolers, P. A. (1976). Reading a year later. *Journal of Experimental Psychology: Human Learning & Memory, 2*(5), 554–565. http://dx.doi.org/10.1037/0278-7393.2.5.554.

Larzabal, C., Tramoni, E., Muratot, S., Thorpe, S. J., & Barbeau, E. J. (2017). Data for: "Extremely long-term memory and familiarity after 12 years." *Mendeley Data, V1*. http://dx.doi.org/10.17632/dzn5wxyxgz.1.

Lawless, H. T. (1978). Recognition of common odors, pictures, and simple shapes. *Perception and Psychophysics, 24*(6), 1973–1975.

Lewis, D. J. (1979). Psychobiology of active and inactive memory. *Psychological Bulletin, 86*(5), 1054–1083.

Light, L. L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42*, 333–376.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*(3), 252–271.

Masquelier, T., & Thorpe, S. J. (2007). Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Computational Biology, 3*(2), e31. http://dx.doi.org/10.1371/journal.pcbi.0030031.

McKone, E., & French, B. (2001). In what sense is implicit memory "episodic"? The effect of reinstating environmental context. *Psychonomic Bulletin & Review, 8*(4), 806–811. http://dx.doi.org/10.3758/BF03196221.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 596–606. http://dx.doi.org/10.1016/S0022-5371(70)80107-4.

Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *The American Journal of Psychology, 112*(1), 113–146.

Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15*(1), 31–49. Retrieved from <http://psycnet.apa.org/journals/xlm/15/1/31.pdf>.

Mitchell, D. B. (2006). Nonconscious priming after 17 years: Invulnerable implicit memory? *Psychological Science, 17*(11), 925–929. http://dx.doi.org/10.1111/j.1467-9280.2006.01805.x.

Mitchell, D. B., & Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(2), 213–222. http://dx.doi.org/10.1037//0278-7393.14.2.213.

Nickerson, R. S. (1968). A note on long-term recognition memory for pictorial material. *Psychonomic Science, 11*(2), 58.

Paivio, A., & Csapo, K. (1973). Picture superiority in free recall: Imagery or dual coding? *Cognitive Psychology, 5*(2), 176–206. http://dx.doi.org/10.1016/0010-0285(73)90032-7.

Roediger, H. L. (2012). Reconsidering implicit memory. In J. Bowers, & C. Marsolek (Eds.). *Rethinking implicit memory* (pp. 3–18). Oxford University http://dx.doi.org/10.1093/acprof:oso/9780192632326.003.0001.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. http://dx.doi.org/10.1111/j.1467-9280.2006.01693.x.

Rubin, D. C., & Schulkind, M. D. (1997). The distribution of autobiographical memories across the lifespan. *Memory & Cognition, 25*(6), 859–866. http://dx.doi.org/10.3758/BF03211330.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103*(4), 734–760.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(3), 501–518. http://dx.doi.org/10.1037/0278-7393.13.3.501.

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior, 6*, 156–163.

Sloman, S. A., Hayman, C. A. G., Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in primed fragment completion. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*(2), 223–239. http://dx.doi.org/10.1037/0278-7393.14.2.223.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50. http://dx.doi.org/10.1037//0096-3445.117.1.34.

Squire, L. R. (1989). On the course of forgetting in very long-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15*(2), 241–245.

Standing, L. (1973). Learning 10000 pictures. *The Quarterly Journal of Experimental Psychology,* (March 2013), 37–41.

Starns, J. J., Chen, T., & Staub, A. (2017). Eye movements in forced-choice recognition: Absolute judgments can preclude relative judgments. *Journal of Memory and Language, 93*, 55–66. http://dx.doi.org/10.1016/j.jml.2016.09.001.

Sutherland, G. R., & McNaughton, B. (2000). Memory trace reactivation in hippocampal and neocortical neuronal ensembles. *Current Opinion in Neurobiology, 10*(2), 180–186. http://dx.doi.org/10.1016/S0959-4388(00)00079-9.

Thorpe, S. J. (2011). Grandmother cells and distributed representations. In N. Kriegeskorte, & G. Kreiman (Eds.). *Visual population codes: Toward a common multivariate framework for cell recording and functional imaging* (pp. 23–51). MIT Press.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373.

Vallet, G. T., Rouleau, I., Benoit, S., Langlois, R., Barbeau, E. J., & Joubert, S. (2016). Alzheimer's disease and memory strength: Gradual decline of memory traces as a function of their strength. *Journal of Clinical and Experimental Neuropsychology, 38*(6), 648–660. http://dx.doi.org/10.1080/13803395.2016.1147530.

Voss, J. L., Baym, C. L., & Paller, K. A. (2008). Accurate forced-choice recognition without awareness of memory retrieval. *Learning & Memory (Cold Spring Harbor, N.Y.), 15*(6), 454–459. http://dx.doi.org/10.1101/lm.971208.

Waters, R. H., & Leeper, R. (1936). The relation of affective tone to the retention of experiences of daily life. *Journal of Experimental Psychology, 19*(2), 203–215. http://dx.doi.org/10.1037/h0062328.

Wixted, J. T. (2009). Remember/Know judgments in cognitive neuroscience: An illustration of the underrepresented point of view. *Learning & Memory, 16*, 406–412. http://dx.doi.org/10.1101/lm.1312809.familiarity.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Review, 2*(6), 409–416.

Yonelinas, A. P. (2001). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General, 130*(3), 361–379. http://dx.doi.org/10.1037//0096-3445.130.3.361.