Something from (almost) nothing: buildup of object memory from forgettable single fixations

Ansgar D. Endress · Mary C. Potter

Abstract We can recognize thousands of individual objects in scores of familiar settings, and yet we see most of them only through occasional glances that are quickly forgotten. How do we come to recognize any of these objects? Here, we show that when objects are presented intermittently for durations of single fixations, the originally fleeting memories become gradually stabilized, such that, after just eight separated fixations, recognition memory after half an hour is as good as during an immediate memory test. However, with still shorter presentation durations, memories take more exposures to stabilize. Our results thus suggest that repeated glances suffice to remember the objects of our environment.

Introduction

We are not only able to identify tens of thousands of objects, from cups to cruisers, but we are also able to remember particular instances of such objects, such as the cup that is on my desk. The objects that populate the many environments we are familiar with seem to be learned effortlessly, can be recognized after incidental exposure (e.g., Castelhano & Henderson, 2005), and yet most may have received no more than an occasional glance. When objects are shown for the duration of a single fixation (about 250 ms; see e.g., Rayner, 1998) or less, those objects are readily processed (e.g., Li, VanRullen, Koch, & Perona, 2002; Potter, 1975), but the resulting memories are fleeting, decaying after a few seconds and disappearing almost completely after a few minutes (e.g., Endress & Potter, 2012; Intraub, 1980; Potter, Staub, Rado, & O'Connor, 2002; Potter, Staub, & O'Connor, 2004). Yet, the same objects are remembered in detail when looked at for 3 s (e.g., Brady, Konkle, Alvarez, & Oliva, 2008).

How do we transform the fleeting memories resulting from single fixations into stable long-term memories? While observers sometimes fixate objects multiple times (e.g., Castelhano, Mack, & Henderson, 2009), they still have to integrate those glimpses into stable LTM representations. Moreover, the time during which an object is projected onto our retina is clearly not the only determinant of the stability or even existence of the resulting memory.

For example, observers famously fail to notice a gorilla in full view when they are attending to other aspects of a display (Drew, Võ, & Wolfe, 2013; Simons & Chabris, 1999) and, at least in statistical learning tasks, observers do not appear to learn anything about those items they do not attend to (Toro, Sinnett, & Soto-Faraco, 2005; Turk-Browne, Jungé, & Scholl, 2005). Conversely, even very brief presentations can yield relatively good memory when viewers attend to them and continue to think about them. For example, presenting pictures for just 110 ms followed by a blank screen for 5,890 ms yields comparable memory

A. D. Endress (✉)
Department of Psychology, City University, Northampton Square, London EC1V 0HB, UK
e-mail: ansgar.endress@m4x.org

A. D. Endress
Universitat Pompeu Fabra, Barcelona, Spain

M. C. Potter
Massachusetts Institute of Technology, Cambridge, MA, USA

to a full 6-s presentation, at least after a retention interval of a few minutes (Intraub, 1980).

If fleeting memories need to be consolidated into more stable ones, there is a second complication, reflected in the long and ongoing history of controversy about the relation between short-term memory (STM) and LTM. The modal model of memory developed in the 1960s held that LTM representations need to be consolidated from STM (Atkinson & Shiffrin, 1968). In contrast, subsequent evidence suggested dissociations between STM and LTM: patients like HM could not form new LTM representations despite relatively spared STM (Scoville & Milner, 1957), and patients like KF could form novel LTM representations despite impaired (verbal) STM (Shallice & Warrington, 1970), leading to the view that STM and LTM might be independent and dissociable (e.g., Baddeley, 2001; Brady, Brady, Konkle, & Alvarez, 2011; Cowan, 2008; Luck, 2007; Sternberg, 2009; see also Craik & Lockhart, 1972, for a compilation of differences between long-term memory and short-term memory).

In fact, there would be an important design reason for a dedicated STM system that is independent of LTM: speed of access. For example, even with memory sets as small as 4–16 items, there is a noticeable delay in accessing memory items that increases with the size of the memory set (Wolfe, 2012). Given the large capacity of LTM (Brady et al., 2008) (that might, in the case of the mental lexicon, comprise more than 50,000 words; see e.g., Pinker, 1999), it is easy to see how access to LTM might be too slow for manipulating mental representations. However, if LTM and STM are indeed independent, it is unclear how we can come to recognize the objects that populate our environments, as we should not be able to transform the fleeting memories resulting from glimpses into stable LTM representations.

While the dominant view is arguably that STM and LTM are independent and dissociable, other authors have proposed STM and LTM to form a continuum (e.g., Crowder, 1993; Greene, 1986; Wickelgren, 1973; Nairne, 2002; see Ranganath, 2005, for a recent review).

Here, we thus ask whether (and how readily) repeated but intermittent exposure to objects for durations of single fixations transforms fleeting STM representations into more stable LTM representations. It is well established since the onset of experimental psychology (Ebbinghaus, 1885/1913) that once an item is in LTM, repeated exposure increases its memory strength and delays its forgetting. It is not clear, however, whether, and how readily, fleeting STM representations would lead to stable LTM representations. In fact, there are reasons to think that brief intermittent exposure might not lead to stable LTM representations. For example, Subramaniam, Biederman, and Madigan (2000) presented participants with rapid sequences of line drawings of objects for a duration of 72 or 126 ms per picture. Prior to each

sequence, participants were given an object name and were instructed to report after the sequence whether that object had appeared in the sequence. Importantly, the items in the sequences were repeated, such that participants had (briefly) seen each picture between 0 and 31 times as nontargets, before that object was used as a target. Subramaniam et al. (2000) found no evidence that detection performance improved as a function of how often the targets had been briefly seen before. Yet when target images were presented for 5 s at the beginning of the experiment (arguably sufficient for LTM), detection performance improved.[1]

At first sight, these results seem to suggest another dissociation between LTM and STM: Repeated exposure clearly improves LTM retention (Ebbinghaus, 1885/1913), while Subramaniam et al.'s (2000) results seem to suggest that this might not be the case for STM. However, in the latter study, participants were actively searching for a target, which, in turn, might have limited how much they remembered of the items, just as observers can fail to remember having viewed a gorilla when they were busy attending to other aspects of a display. Perhaps if participants were simply viewing items without a search task, memory would improve with repetition.

In fact, there are other reasons to think that STM might be consolidated across views. For example, Pertzov, Avidan, and Zohary (2009) presented participants with an array of objects, and measured short-term retention of the objects as a function of how often they had been fixated. Results showed that objects were remembered better when they had been fixated more often. However, while these results suggest that, within STM, memory can be strengthened by multiple views, they leave open the question of whether such strengthening would occur over longer retention intervals as well. Likewise, Melcher (2001) found that memory for objects improved if participants were shown a scene containing many objects repeatedly for 0.25, 1, or 2 s, and if they had to recall objects from the scene after each presentation. He proposed that the memory was an intermediate form of memory, between STM and LTM (see also Melcher & Kowler, 2001; Melcher, 2006). However, these results leave open the question of whether memory consolidation would also occur with the shortest presentation durations if the presentations were not followed by a recall period;

---

[1]In a different condition, Subramaniam et al. (2000) showed that priming occurred even with a 72-ms presentation, provided that the prime was designated as the target on that trial and was followed by a 3-s interval before the sequence began. However, and as mentioned above, Intraub (1980) showed that presenting an item for 110 ms followed by a 5,890-ms blank screen yields comparable memory performance to presenting the item for 6,000 ms. In line with these results, we thus propose that the priming Subramaniam et al. (2000) observed under these conditions might not have been due to the brief presentation per se, but rather to the time available to process the stimuli during the presentation and the following blank interval.

after all, this period gives participants additional processing time even after the retinal image had disappeared, and it is well known that repeated recall leads to improved memory even when the memory items are presented only once (e.g., Erdelyi & Becker, 1974).

Here, we investigate directly whether repeated brief exposure to objects can transform fleeting STM representations into more stable LTM representations. Experiment 1 comprised two parts. In each trial of the first part, participants were presented with a sequence of 12 pictures of everyday objects for a duration of 250 ms per picture. A test picture followed: participants had to decide whether it had been part of the sequence. Importantly, most pictures were shown only once throughout the experiment, but a subset appeared in 1, 2, 4, 8, or 16 trials, with repetitions separated by three to seven other trials (each with 12 pictures), equivalent to an average delay of about 50 s. The repeated items were never used as test items in the first part.

In the second part of the experiment, we tested participants' LTM for these repeated pictures (including the subset that was presented only once and not repeated). Experiment 2 was similar to experiment 1, except that items were repeated only for up to eight times, and that we also tested the effect of the retention interval between the last presentation of an item and the later LTM test.

To foreshadow our results, experiments 1 and 2 revealed better long-term retention when items were repeated more often. In experiment 3, we tested two critical issues arising from these results. First, other studies using shorter presentation durations did not find memory consolidation; we thus asked whether items presented more briefly would yield a memory benefit, and whether this benefit was reduced compared to longer presentation durations. Second, we asked whether long-term memory retention was simply a function of the total presentation time.

Finally, in experiment 4, we compared the behavior with pictorial stimuli from experiments 1 to 3 to that with verbal stimuli, for two important reasons. First, there is substantial evidence that (short-term) memory for verbal items is independent of other forms of memory (e.g., Baddeley, 1996, 2003; Endress & Potter, 2012). Further, while repeatedly recalling pictorial stimuli improves recall performance for the memory items, this does not seem to be the case for words (e.g., Erdelyi & Becker, 1974). Hence, it is possible that words would behave differently from pictures in the experiments outlined above. Second, and crucially, the problem participants faced in experiment 4 is different from that in experiments 1 to 3. In experiments 1 to 3, participants had to construct novel LTM representations, and to decide later on whether or not they had *ever* seen the test items. In experiment 4, in contrast, participants had pre-existing LTM representations for the words they were presented with (otherwise they would have been non-words). Hence, during

the LTM test, participants had to distinguish between words they knew and saw during the experiment, and words they knew but had not seen during the experiment. As a result, it is possible that this latter task might be somewhat more difficult than that in experiments 1 to 3.

General methods

Participants

We recruited 48 participants (nine, eight, and eight females in experiments 1 to 3, mean age 24.2) from the MIT community, sequentially assigned to experiments 1 to 3. An additional 18 participants (11 females, mean age 22.2) were recruited for experiment 4; we retained only the first 16 for analysis to keep the number of participants constant across experiments.

Apparatus

Stimuli were presented on a NEC MultiSync FE700+ 17" CRT (refresh rate: 75 Hz; resolution: 640 × 480), using the Matlab psychophysics toolbox (Version 3.0.8; Brainard, 1997; Pelli, 1997). Responses were collected from pre-marked "Yes" and "No" keys on the keyboard.

Materials

Stimuli in experiments 1 to 3 were color photographs of familiar everyday objects taken from Brady et al. (2008). Stimuli were randomly selected for each participant from a set of 2,400 pictures. They were presented subtending a visual angle of 12.7 × 12.7 degrees.

Stimuli in experiment 4 were words. We selected 2,381 nouns from the CELEX database (a frequency database for English words) with the constraints that each noun (i) had between four and ten letters; (ii) had one or two syllables; (iii) had a minimum frequency of 100 of out 17.9 million words; (iv) was unique in the final list (e.g., words that differed only in plural markers were removed); (v) was not specific to British English; (vi) was not a proper noun; (vii) was not a swear word or otherwise offensive. Words were presented in a font size of 22 in Courier lowercase font.

Experiment 1

Methods

In the first part of the experiment, participants started each trial by a key-press; after a fixation cross, 12 pictures were

presented for a duration of 250 ms per picture. 1.5 s after the sequence the test picture was presented for 800 ms. "Old" test items (those that occurred in the sample sequence) were randomly sampled from two initial, two final and two middle positions, excluding the first and last pictures. "New" test items were presented on 50 % of the trials.

There were 50 critical pictures that were presented in 1, 2, 4, 8, or 16 trials each (hereafter called the "repeated pictures" even though ten were presented only once). None of the repeated pictures were tested immediately after the sequence, and participants were not informed that pictures would be repeated or that there would be a later memory test of some of the items. There was a total of 132 trials. Trials were organized into five start trials, 30 end trials, and 97 central trials. The repeated pictures were presented only in the central trials. Picture repetitions were separated by five intervening trials on average (minimum: three; maximum seven), corresponding on average to a delay of about 50 s. Serial position in the sequence was counterbalanced within each number of repetitions; the repeated pictures were never in the first or last serial position.

In the second part of experiment 1, participants were tested on their memory of the repeated pictures. They were shown all 50 repeated pictures (10 items × 5 numbers of repetitions) mixed with 50 novel pictures, presented one at a time. For each picture, participants indicated whether or not they had seen it before in the experiment. We unconfounded the number of repetitions and the memory delay by controlling the last appearance of the memory items during the first phase of the experiment. In the first phase, the last presentation of all critical items was within a range of six trials. As a result, the delay between the last presentation of a given picture and when it was tested in the second phase was unrelated to the number of repetitions.

Results and discussion

In the immediate memory test of the first part of experiment 1, participants performed well above chance (M = 75.8 %, SD = 4.9 %; Fig. 1), $t(15) = 21.2$, $p < .0001$, Cohen's $d = 5.3$. Using the two-high threshold formula for estimating memory capacity (e.g., Cowan, 2001; Rouder et al., 2008), this performance corresponds to a capacity of 6.2.

In the LTM test in the second part, participants were not above chance when the picture had been presented only once (M = 54.1 %; SD = 9.0 %; Fig. 1 and Table 1), suggesting that a single fixation is rarely sufficient to establish a stable long-term memory.

However, when the critical items were repeated, long-term retention improved such that, after eight repetitions, participants' LTM performance was equivalent to that in the immediate tests in the first part, and after 16 repetitions, LTM performance exceeded performance on the immediate
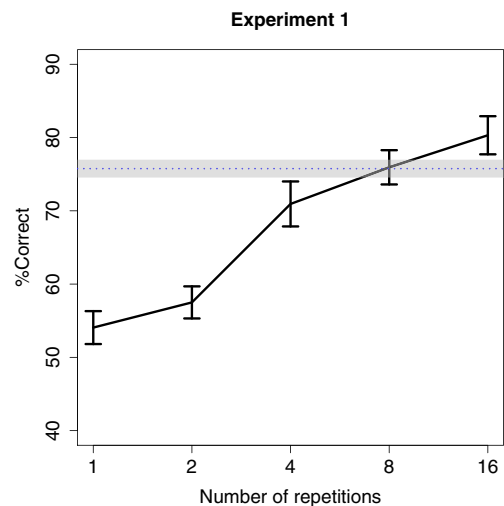

Fig. 1 Results of experiment 1. The *dotted line* shows the average percentage of correct responses during the first part of experiment 1 with an immediate test. The *shaded region* around it represents the associated SEM. The *solid line* shows the percentage of correct responses in the second part of experiment 1 as a function of the number of repetitions. *Error bars* represent SEM

memory test (see Table 1). The LTM results were analyzed using a logistic mixed-effects model with the number of repetitions as a linear predictor. We observed an effect of the number of repetitions, $\beta = .085$, $Z = 7.6$, $p < .0001$, suggesting that LTM retention improved when items were repeated more often.

Experiment 2

Method

In experiment 2, we investigated the temporal stability of newly created LTM representations. experiment 2 was similar to experiment 1, with two exceptions. First, due to the limited set of available pictures, items were repeated only up to eight times. Second, and crucially, we included a manipulation of the retention delay. Specifically, as in experiment 1, the first part of experiment 2 measured short-term retention of the items. This part included 156 trials that were organized into five start trials, four end trials, and three central blocks of 49 trials each. Each central block contained 40 critical pictures that were presented in one, two, four, or eight trials each, and whose long-term retention would be tested in the second part. Constraints on the number of intervening trials between repetitions of a picture as well as on the serial positions of the repeated pictures were as in experiment 1.

The central blocks were used to manipulate the retention interval by inverting the order of appearance of the blocks between parts 1 and 2. Specifically, in part 2, long-term

Table 1 Percentage of correct responses and associated *t* tests as a function of the number of repetitions in the second part of each experiment and of block (Experiments 2 and 4 only; block 3 corresponds to the shortest retention delay while block 1 corresponds to the longest one)

Block	# Repet.	M	SD	Against chance level of 50 %			Against 1st part performance		
				t(15)	p	Cohen's d	t(15)	p	Cohen's d
Experiment 1 (second part)									
–	1	54.1	9.0	1.8	.091	.45	10.4	<.0001	2.6
–	2	57.5	8.8	3.4	.0038	.86	6.6	<.0001	1.7
–	4	70.9	12.3	6.8	<.0001	1.7	1.5	.155	.4
–	8	75.9	9.3	11.1	<.0001	2.8	.1	.941	.02
–	16	80.3	10.4	11.6	<.0001	2.9	2.3	.038	.57
Experiment 2 (second part)									
3	1	55.6	11.2	2.0	.064	.5	7.2	<.0001	1.8
3	2	61.9	14.4	3.3	.005	.83	4.3	.0007	1.1
3	4	67.2	15.2	4.5	.0004	1.1	2.3	.035	.58
3	8	73.4	13.3	7.1	<.0001	1.8	.6	.543	.16
2	1	56.6	9.6	2.7	.016	.68	7.3	<.0001	1.8
2	2	57.5	9.5	3.2	.006	.79	6.5	.0001	1.6
2	4	65.0	15.8	3.8	.002	.95	3.1	.008	.76
2	8	72.8	11.5	7.9	<.0001	2.0	.8	.415	.21
1	1	54.1	12.0	1.3	.196	.34	7.0	<.0001	1.7
1	2	52.8	8.6	1.3	.208	.33	9.9	<.0001	2.5
1	4	61.2	12.0	3.7	.002	.93	4.0	.001	1.0
1	8	70.9	10.2	8.2	<.0001	2.1	1.8	.088	.46
Experiment 3 (second part)									
–	1	49.4	6.8	.4	.718	.10	9.6	<.0001	2.4
–	3	55.6	8.3	2.7	.017	.67	4.4	.0005	1.1
–	6	54.7	10.7	1.8	.101	.44	4.7	.0003	1.2
–	12	60.9	10.4	4.2	.0007	1.1	2.5	.025	.62
–	24	61.6	13.1	3.5	.003	.88	2.1	.053	.53
Experiment 4 (second part)									
3	1	53.75	11.0	1.36	0.194	0.34	5.8	<.0001	1.5
3	2	54.69	12.3	1.52	0.149	0.38	6.18	<.0001	1.5
3	4	55	10.5	1.91	0.076	0.48	6.41	<.0001	1.6
3	8	63.44	13.7	3.91	0.001	0.98	4.02	0.001	1.0
2	1	55	10.6	1.88	0.08	0.47	6.94	<.0001	1.7
2	2	53.12	5.7	2.18	0.0457	0.54	7.11	<.0001	1.8
2	4	53.12	10.5	1.19	0.251	0.3	7.28	<.0001	1.8
2	8	53.44	10	1.38	0.187	0.35	6.12	<.0001	1.5
1	1	53.44	9.6	1.43	0.173	0.36	7.04	<.0001	1.8
1	2	52.5	13.2	0.76	0.459	0.19	5.77	<.0001	1.4
1	4	56.56	12.9	2.04	0.06	0.51	4.35	0.0006	1.1
1	8	59.06	12.9	2.8	0.01	0.7	4.74	0.0003	1.2

The first *t* test is evaluated against the chance level of 50 %. The second *t* test is a paired *t* test comparing the performance in the first part to the performance in the second part for the different blocks and numbers of repetitions

retention was tested first for items from block 3, and last for items from block 1. As a result, the retention interval was longest for items from block 1 and shortest for items from block 3, making the variable presentation-block an approximation of the retention interval. The average delay between the most recent presentation and the test was 3 min 50 s

(SE = 11 s) for block 3, 16 min 19 s (SE = 1 min 6 s) for block 2, and 29 min 55 s (SE = 2 min 54 s) for block 1. The blocks were presented in a continuous sequence of trials, without breaks.

The second part of experiment 2 comprised 240 test pictures (120 repeated pictures, 120 novel foils).

Results

In the immediate memory test in the first part of experiment 2, participants again performed well above chance (M = 75.3 %, SD = 7.3 %; Fig. 2), $t = 13.8$, $p < .0001$, Cohen's $d = 3.5$. Using the two-high threshold formula for estimating memory capacity (e.g., Cowan, 2001; Rouder et al., 2008), this performance corresponds to a capacity of 6.1.

In the LTM test in the second part, participants performed at or close to chance when the picture had been presented only once (M = 55.6 %, SD = 11.2 %; Fig. 2 and Table 1), suggesting again that a single fixation is rarely sufficient to establish a stable long-term memory.

However, when the critical items were repeated, long-term retention improved such that, after eight repetitions, participants' LTM performance was indistinguishable from the immediate tests in the first part (see Table 1). The LTM results were analyzed using a logistic mixed-effects model

with linear predictors of block and of number of repetitions as well as their interaction. Only the two main effects contributed to the model likelihood. As in experiment 1, we observed a significant main effect of number of repetitions, $\beta = .11$, $Z = 8.5$, $p < .0001$, suggesting again that LTM retention improved when items were repeated more often. Further, we observed a significant main effect of block (i.e., our proxy for the retention interval), $\beta = .11$, $Z = 2.5$, $p = .011$, suggesting that LTM retention decreased with increasing memory delay, as seen in Fig. 2b.

Discussion

The results of experiment 2 yielded two crucial results. First, repeating items more often allows participants to consolidate fleeting STM traces into more durable LTM representations. When items were presented only once, virtually no long-term retention was observed; when they were presented eight times, long-term retention performance was indistinguishable from an immediate memory test. This is not to say that a single presentation might not yield memory traces that might be detectable by more sensitive methods. In fact, if the fleeting memories disappeared immediately, it is hard to see how additional exposure could possibly help to stabilize them. Rather, these results show that brief exposure to an item yields memory traces that diminish in strength so rapidly that they cannot be detected by a recognition task a few minutes later; this temporal dynamic is arguably a quality that is usually associated with STM. Crucially, additional exposure stabilizes these fleeting memories so that they can be detected about half an hour later, which, in turn, is a quality associated with LTM.

The second crucial result of experiments 1 and 2 is that long-term retention performance declined with increased retention intervals. This raises the question of whether the performance decrement was due to decay or interference, or both. In fact, interference might well play a crucial role in long-term forgetting (e.g., Brown, Neath, & Chater, 2007; see also Melcher & Murphy, 2011), which, in turn, does not exclude the possibility that decay might also be a factor in forgetting (e.g., Hollingworth, 2005). While the current experiments do not allow us to decide whether the reduction in long-term retention after longer intervals was due to decay, interference or both, experiments using the same stimuli as those employed here suggest that, at least in the short-term domain, decay might play a role. Specifically, in Endress and Potter's (2014) experiment 4, participants were presented with a task similar to part 1 of experiments 1 and 2. Crucially, participants were tested either after a 1.5-s retention interval or after a 7.5-s retention interval. Participants performed worse after the 7.5-s delay than after the 1.5-s delay. As the only difference between the delay conditions was the duration of a progress bar participants viewed
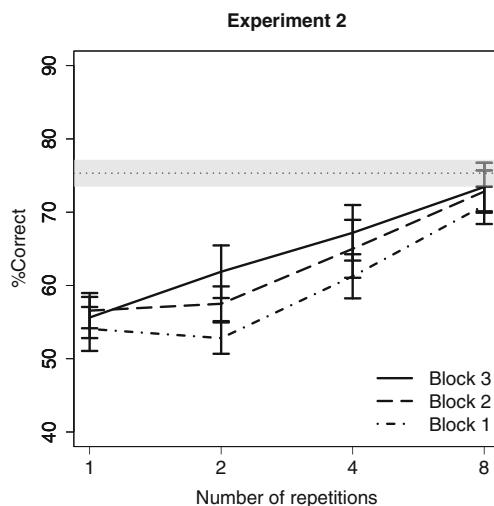

Experiment 2

Fig. 2 Results of experiment 2. The *dotted line* shows the average percentage of correct responses during the first part of experiment 2 with the immediate test. The *shaded region* around it represents the associated SEM. The other lines show the percentage of correct responses in the second part of experiment 2 as a function of block and the number of repetitions. The block number is given relative to the first part of the experiment, and the order of blocks was inverted in the second part. The *solid line* represents the performance in block 3 (average delay between the most recent presentation and the test, 3 min 50 s, SE = 11 s), the *dashed line* the performance in block 2 (average delay 16 min 19 s, SE = 1 min 6 s), and the *dash-dotted line* the performance in block 1 (average delay 29 min 55 s, SE = 2 min 54 s). *Error bars* represent SEM

between the end of the sample sequence and the test item, it is plausible to attribute this difference to decay. If these results scale up to long-term retention, one would expect a role of memory decay in long-term retention as well, in conjunction with interference from viewing many other images before seeing the test images.

However, the combined results of experiments 1 and 2 raise two crucial questions. First, given that a prior study with shorter presentation durations (Subramaniam et al., 2000) found no priming benefit for repeated items, we asked whether we would observe any memory consolidation with briefer presentations. Second, we asked whether the total viewing time would predict long-term retention, or whether it would be modulated by the presentation duration. We thus presented items for 133 ms each, but repeated them up to 24 times. As a result, for the largest number of repetitions, both the total viewing time and the number of repetitions was larger than in experiments 1 and 2.

Experiment 3

Methods

The design of experiment 3 was based on that of experiment 2, except that the same items were repeated across blocks. As a result, the critical items were presented 1, 3, 6, 12, or 24 times. Further, we included an additional 24 end trials (28 in total), so that the long-term retention test would not start immediately after the last presentation of the memory items. Crucially, items were presented for 133 ms per picture rather than 250 ms.

Results

Results of experiment 3 In the immediate memory test of the first part of experiment 3, participants performed well above chance ($M = 67.7\,\%$, $SD = 7.1\,\%$; Fig. 3), $t(15) = 10.0$, $p < .0001$, Cohen's $d = 2.5$, $CI_{.95} = 64.0, 71.5$. Using the two-high threshold formula for estimating memory capacity (e.g., Cowan, 2001; Rouder et al., 2008), this performance corresponds to a capacity of 4.2.

In the LTM test in the second part, participants were not above chance when the picture had been presented only once (M = 49.4 %; SD = 6.8 %; Fig. 3 and Table 1), suggesting that a single fixation is rarely sufficient to establish a stable long-term memory.

However, when the critical items were repeated, long-term retention gradually improved. In contrast to experiments 1 and 2, however, LTM performance never reached the level of the immediate memory test (see Table 1).
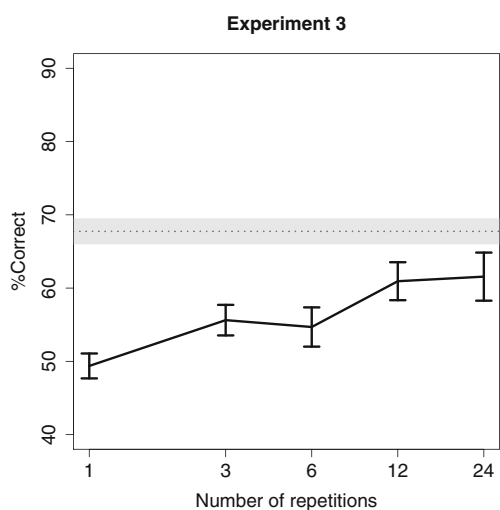

Fig. 3 Results of experiment 3. The *dotted line* shows the average percentage of correct responses during the first part of experiment 3 with an immediate test. The *shaded region* around it represents the associated SEM. The *solid line* shows the percentage of correct responses in the second part of experiment 3 as a function of the number of repetitions. *Error bars* represent SEM

The LTM results were analyzed using a logistic mixed-effects model with the number of repetitions as a linear predictor. We observed an effect of the number of repetitions, $\beta = .019$, $Z = 3.1$, $p = .002$, suggesting that LTM retention improved when items were repeated more often.

Comparison with experiment 1 We compared experiment 3 to experiment 1 for three reasons. First, both experiments comprised only a single block of items. Second, both experiments had a similar number of end trials (30 in experiment 1 vs. 28 in experiment 3), making the memory delay roughly comparable, although it was slightly shorter in experiment 3, due to both the smaller number of end trials and to the shorter presentation duration. Third, the maximal number of repetitions as well as the maximal total viewing duration were greater in experiment 3 than in experiment 1.

We compared experiments 1 and 3 in two ways. First, we compared performance on the immediate retention task. Participants in experiment 1 ($M = 75.8\,\%$) performed better than in experiment 3 ($M = 67.7\,\%$), $t(30) = 3.73$, $p = .0008$, Cohen's $d = 1.32$.

Second, we asked whether the speed of learning was different across experiments 1 and 3, and whether LTM retention differed across the experiments. The data points we used as a proxy for long-term retention were the performance for items repeated eight times in experiment 1 ($M = 75.9\,\%$, $SD = 9.3\,\%$), and for those repeated 24 times in experiment 3 ($M = 61.6\,\%$, $SD = 13.1\,\%$). As a result, the items from experiment 3 were presented both more often than in experiment 1 (24 vs. 8 times) and for a longer total presentation duration (3.2 s vs. 2 s). Hence, if we observe

an impairment in long-term retention performance with the shorter presentation duration, this impairment occurred despite more repetitions and a longer total duration. We thus submitted the data to an ANOVA with the between-subjects predictor presentation duration (i.e., experiment 1 vs. 3) and the within-subject predictor number of repetitions (i.e., 1 vs. 8 repetitions for experiment 1 and 1 vs. 24 repetitions for experiment 3).

The analysis revealed a main effect of presentation duration, $F(1, 30) = 11.2$, $p = .002$, $\eta_p^2 = .271$, as well as of the number of repetitions, $F(1, 30) = 73.4$, $p < .0001$, $\eta_p^2 = .671$. Crucially, the two factors interacted, $F(1, 30) = 5.9$, $p = .021$, $\eta_p^2 = .0543$.

This interaction can be interpreted in two ways. First, separate ANOVAs for the different numbers of repetitions (once vs. more than once) revealed no difference between experiments 1 and 3 when items were presented only once, $F(1, 30) = 2.8$, $p = .107$, $\eta_p^2 = .08$, but a sizable difference when items were repeated more than once, $F(1, 30) = 12.7$, $p = .001$, $\eta_p^2 = .298$. Hence, even though total viewing time and number of repetitions were both greater in experiment 3, long-term memory performance was worse.

Second, separate ANOVAs for the two presentation durations with the within-subject factor number of repetitions revealed that the difference between performance for items presented only once and items presented more than once was larger in experiment 1 (average difference: 21.9 %, $SE = 3.0$ %), $F(1, 15) = 54.9$, $p < .0001$, $\eta_p^2 = .785$, than in experiment 3 (average difference: 12.2 %, $SE = 2.7$ %), $F(1, 15) = 21.0$, $p = .0004$, $\eta_p^2 = .583$, suggesting that LTM memory consolidation was better in experiment 1 than in experiment 3 even though total viewing time and number of repetitions were both greater in experiment 3.

Discussion

Experiment 3 yielded three crucial results. First, and as in experiments 1 and 2, long-term retention was poor or non-existent when items were presented once, and improved when items were presented more often. Hence, experiment 3 further supports the idea that fleeting memories can be transformed into more stable memory representations simply by presenting items more often.

Second, both the absolute level of long-term retention and the benefit for each additional repetition were reduced for shorter presentation durations. These results suggest that the stabilization of memories is graded and depends on how strong the memory representations were in the first place. Hence, it is possible that, with still shorter presentation durations, no long-term retention at all might be observed.

Third, our results show that the total time during which an item has been viewed is a poor predictor of long-term retention, although it can be a good predictor with somewhat longer durations or extra processing time (e.g., Melcher, 2001). As mentioned above, long-term retention after 24 repetitions in experiment 3 was substantially worse than long-term retention after eight repetitions in experiment 1, although both the number of repetitions and the total viewing durations were greater. Hence, retention is not a monotonic function of the time during which an item has been viewed, suggesting that there might be some "leakage" when integrating very weak memory representations over time. These results thus further support the idea that the benefit of each additional repetition of a stimulus depends on the strength with which the stimulus has been represented in the first place.

Experiment 4

In experiment 4, we further extend the results of experiments 1 to 3 by asking whether similar results can be found with verbal stimuli. As mentioned before, the objective of this experiment was (i) to establish that intermittent repetitions of verbal stimuli improve LTM retention, and (ii) to test whether pre-existing LTM representations might impair recognition memory.

Method

Experiment 4 was similar to experiment 2, with two exceptions. First, we used words rather than pictures as stimuli. Second, stimuli were presented for a duration of 147 ms per word, as pilot experiments revealed that this presentation duration yielded a comparable short-term level of retention to that of experiment 2.

Results and discussion

In the immediate memory test in the first part of experiment 4, participants again performed well above chance ($M = 74.0$ %, $SD = 8.6$ %; Fig. 4), $t = 11.2$, $p < .0001$, Cohen's $d = 2.8$. This result was similar to the part 1 results of experiment 2. In the LTM test in the second part, participants performed at or close to chance when the word had been presented only once ($M = 53.4$ %, $SD = 9.6$ %; Fig. 4 and Table 1), suggesting again that a single fixation is not sufficient for recognizing items after a retention time of a few minutes.

When the critical items were repeated, LTM retention gradually improved, and differed from chance performance when items were repeated eight times. However, LTM per-
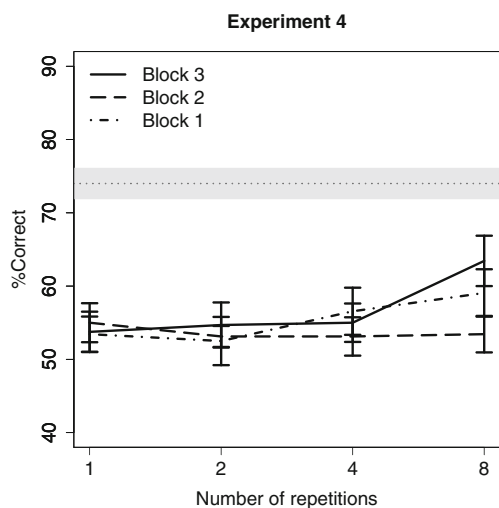
Experiment 4

Fig. 4 Results of experiment 4. The *dotted line* shows the average percentage of correct responses during the first part of experiment 4 with an immediate test. The *shaded region* around it represents the associated SEM. The *solid line* represents the long-term retention performance in block 3 (average delay between the most recent presentation and the test, 3 min 9 s, SE = 7 s), the *dashed line* the performance in block 2 (average delay 12 min 27 s, SE = 19 s), and the *dash-dotted line* the performance in block 1 (average delay 22 min 17 s, SE = 39 s). *Error bars* represent SEM

formance remained substantially worse than performance in the short-term retention task even with eight repetitions (see Table 1).

The LTM results were analyzed using a logistic mixed-effects model with linear predictors of block and of number of repetitions as well as their interactions. Only the two main effects contributed to the model likelihood. We observed a significant main effect of the number of repetitions, $\beta = .029$, $Z = 2.41$, $p = .016$, suggesting again that LTM retention improved when items were repeated more often. The main effect of block did not reach significance, $\beta = .027$, $Z = .68$, $p = .50$, ns.

The combined LTM results of experiments 2 and 4 were analyzed using a logistic mixed-effects model with linear predictors of experiment, block, and of number of repetitions as well as their interactions. We observed significant main effects of the number of repetitions, $\beta = .029$ $Z = 2.67$, $p = .0075$, and of block, $\beta = .064$, $Z = 2.24$, $p = .025$. We also observed an interaction between the number of repetitions and the experiment, $\beta = .082$ $Z = 6.02$, $p < .0001$, suggesting that LTM performance increased more as a function of the number of repetitions in experiment 2 than in experiment 4.

As mentioned above, the reduced benefit of repeating words as compared to pictures might be due to the fact that all words were known before the experiment; hence, participants needed to discriminate known items that have been seen recently from known items that have not, while participants in experiments 1 to 3 just needed to discriminate known items from unknown items. This, in turn, might have made the LTM retention task in experiments 1 and 2 easier.

General discussion

The present results reveal that repeatedly seeing objects for the duration of a single fixation transforms fleeting STMs into more stable LTM traces. After seeing an object once for the duration of a single fixation, almost no LTM retention was observed. Further, previous research has shown that these fleeting memories start decaying within seconds (Endress & Potter, 2012, 2014; Intraub, 1980; Potter et al., 2002, 2004), suggesting that they are firmly in the domain of STM.

In contrast, after seeing an object for just eight 250-ms durations (with each presentation separated by about 1 min and the viewing of about 60 other pictures), recognition performance 30 min later was indistinguishable from an immediate test, and in experiment 1, LTM recognition performance exceeded an immediate memory test after 16 repetitions. However, when the presentation duration was reduced to 133 ms per picture, LTM built up more slowly. We suggest that repeatedly viewing objects in the environment, even for durations of single fixations, gradually builds up LTM representations, and allows us to recognize the tens of thousands of objects we know. The speed of the built-up might depend on various factors, including the strength of the initial representation, whether or not items are attended, and other factors. Moreover, the results of experiment 2 show that these initial LTM representations show some degree of loss over tens of minutes, a decline that might plausibly continue over longer retention intervals (Nairne, 1992). Our and previous results lend support to the idea that there is a continuum between STM and LTM (see Ranganath & Blumenfeld, 2005 for a recent review). In fact, the best-accepted dissociations between LTM and STM are that "only short-term memory [demonstrates] (1) temporal decay and (2) chunk capacity limits" (e.g., Cowan, 2008). However, both dissociations might need to be revised. First, while LTM is thought to be relatively stable over time, the current and previous results (e.g., Nairne, 1992) show that there is still temporal decay. Second, the capacity of STM might actually be larger than has been assumed, and the capacity limitations traditionally attributed to STM might be an effect of proactive interference. Specifically, Endress and Potter (2014) presented participants with rapid sequences of pictures or words, and showed that participants briefly retain a certain proportion of the presented items rather than a fixed number of items (e.g., about 30 pictures out of

100 pictures), and that this proportion is relatively independent of the number of presented items. The limited memory capacity that is thought to be characteristic of STM arose only under conditions of strong proactive interference.[2] Hence, in contrast to the widely held view about differences between STM and LTM, LTM might undergo decay, and STM might have a large capacity. Together with the current ones, such results thus raise the possibility that STM and LTM might form a continuum, with short-lasting memories for brief stimuli constituting a fragile form of long-term memory.

However, this continuum might take one of two forms. First, STM and LTM might reflect the same psychological construct. Second, STM and LTM representations might be distinct but be created in parallel, with the initial stages of LTM representations having very similar properties to those usually attributed to STM: an LTM representation might be generated by a brief exposure, and might be initially unstable, decaying over the course of a few seconds or minutes.[3] The relative temporal stability that is the hallmark of LTM items might be achieved only after repeated or prolonged exposure. Further, the limited capacity that is thought to be a hallmark of temporary memory systems (e.g., Cowan, 2008) might be observable only under conditions of strong proactive interference; otherwise, temporary memory might have an open-ended capacity similar to LTM (Endress & Potter, 2014). As a result, if STM and LTM are independent, the initial stages of LTM must have properties that are characteristic both of LTM (e.g., a large capacity) and of STM (e.g., availability after brief exposure; susceptibility to decay or interference).

Be that as it might, our results suggest that brief exposure to objects results in unstable LTM memory traces, either because STM and LTM reflect the same construct, or because the initial stages of LTM representations have similar properties to STM representations. Once an initial

memory trace is created, repeated but intermittent fixations of objects might be sufficient to establish a stable representation of the objects in each of the thousands of contexts with which we are familiar, building up a detailed knowledge base that enables us to operate efficiently in the complex environment we inhabit.

References

Atkinson, R.C., & Shiffrin, R.M. (1968). In K. W. Spence & J. T. Spence (Eds.), *Human memory: A proposed system and its control processes* (Vol. 2, pp. 89–195). New York: Academic Press.

Baddeley, A. D. (1996). The fractionation of working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(24), 13468–13472.

Baddeley, A. D. (2001). Is working memory still working. *American Psychologist*, *56*(11), 851–864.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, *11*(5), 4.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(38), 14325–14329.

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576.

Castelhano, M. S., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition*, *12*(6), 1017–1040.

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*, 3.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, *169*, 323–338.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671–684.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory and Cognition*, *2*, 142–145.

Drew, T., Võ, M. L.-H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattentional blindness in expert observers. *Psychological Science*, *24*(9), 1848–1853.

Ebbinghaus, H. (1885/1913). *Memory: A contribution to experimental psychology*. New York: Teachers College, Columbia University. http://psychclassics.yorku.ca/Ebbinghaus/

---

[2] Endress and Potter (2014) manipulated proactive interference by presenting participants either with objects that would appear only in a single trial, or by using the same set of objects throughout all trials. When objects are trial-unique, proactive interference should be minimized. In contrast, when objects from the same set of items are used across trials, there should be interference from previous trials, because participants would have to decide whether they have seen a given test item in the current trial, or rather in one of the previous trials. Accordingly, Endress and Potter (2014) found large memory capacity estimates under low-interference conditions, while capacity estimate remained in the range of previous working memory estimates in the high-interference condition. Delayed recognition tests showed that these initially remembered items began to be forgotten within seconds, and almost all were forgotten within 30 min.

[3] This is not to imply that a single, brief presentation might not yield some weak LTM trace; rather our results suggest that such traces are generally too weak to be detect in a simple recognition paradigm.

Endress, A. D., & Potter, M. C. (2012). Early conceptual and linguistic processes operate in independent channels. *Psychological Science*, *23*(3), 235–245.

Endress, A. D., & Potter, M. C. (2014). Large capacity temporary visual memory. *Journal of Experimental Psychology. General*, *143*(1), 548–565.

Erdelyi, M. H., & Becker, J. (1974). Hypermnesia for pictures: Incremental memory for pictures but not words in multiple recall trials. *Cognitive Psychology*, *6*(1), 159–171.

Greene, R. L. (1986). A common basis for recency effects in immediate and delayed recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *12*(3), 413.

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(3), 396–411.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(1), 1–12.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9596–9601.

Luck, S. J. (2007). Visual short term memory. *Scholarpedia*, *2*(6), 3328.

Melcher, D. (2001). Persistence of visual memory for scenes. *Nature*, *412*(6845), 401.

Melcher, D. (2006). Accumulation and persistence of memory for natural scenes. *Journal of Vision*, *6*(1), 8–17.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research*, *41*(25–26), 3597–3611.

Melcher, D., & Murphy, B. (2011). The role of semantic interference in limiting memory for the details of visual scenes. *Frontiers in Psychology*, *2*, 262.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science*, *3*(3), 199–202.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, *53*, 53–81.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pertzov, Y., Avidan, G., Zohary, E. (2009). Accumulation of visual information across multiple fixations. *Journal of Vision*, *9*(10), 2.1–212.

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York: Basic Books.

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966.

Potter, M. C., Staub, A., & O'Connor, D. H. (2004). Pictorial and conceptual representation of glimpsed pictures. *Journal of Experimental Psychology. Human Perception and Performance*, **30**(3), 478–489.

Potter, M. C., Staub, A., Rado, J., & O'Connor, D. H. (2002). Recognition memory for briefly presented pictures: The time course of rapid forgetting. *Journal of Experimental Psychology. Human Perception and Performance*, *28*, 1163–1175.

Ranganath, C., & Blumenfeld, R. S. (2005). Doubts about double dissociations between short- and long-term memory. *Trends in Cognitive Sciences*, *9*(8), 374–380.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 5975–5979.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry*, *20*(1), 11–21.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly Journal of Experimental Psychology*, *22*(2), 261–273.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074.

Sternberg, R. J. (2009). *Cognitive psychology* (5th ed.). Belmont: Wadsworth.

Subramaniam, S., Biederman, I., & Madigan, S. (2000). Accurate identification but no priming and chance recognition memory for pictures in rsvp sequences. *Visual Cognition*, *7*(4), 511–535.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), B25–B34.

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology. General*, *134*(4), 552–564.

Wickelgren, W. A. (1973). The long and the short of memory. *Psychological Bulletin*, *80*(6), 425–438.

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, *23*(7), 698–703.