Time-Dependent Forgetting and Retrieval Practice Effects in Detailed Visual Long-Term Memory

Tom Mercer and Gemma A. Jones

University of Wolverhampton

Correspondence concerning this article should be addressed to Tom Mercer, Faculty of Education, Health and Wellbeing, University of Wolverhampton, Wolverhampton, UK, WV1 1LY. Email: t.mercer2@wlv.ac.uk. Telephone: +44(0)1902 321368.

Acknowledgements

We are grateful to Talia Konkle and Denis McKeown for permission to use their stimuli. We are also indebted to Julie Jebsen for her assistance with data scoring in Experiment 3.

Disclosure statement

The authors report no conflicts of interest.

ABSTRACT

Memories – especially those containing fine details – are usually lost over time, but the present study assessed whether detailed visual memories can survive a one-week delay if retrieval practice is provided. In three experiments, participants viewed 300 objects and then completed recognition tests assessing memory for precise object exemplars and their state. The recognition tests occurred immediately after encoding and one week later, and required participants to distinguish between a previously seen target object and an incorrect foil. Whilst there was forgetting when participants were tested on different sets of stimuli across the delay, retrieval practice led to an advantage in recognition performance. This effect was not simply due to mere exposure, as retrieval practice boosted recognition beyond a restudy condition, which had a second encoding opportunity but no retrieval practice. Yet more detailed analyses revealed that the effect of retrieval practice was highly dependent upon the type of information being tested (exemplar or state) and the specific foil that was presented. In addition, state information was harder to retain over the delay than exemplar information, suggesting that memory for different properties is forgotten at different rates.

Key words: forgetting, visual long-term memory, time, retrieval practice.

Memories seem to be forgotten over time. This form of forgetting was empirically demonstrated in the pioneering work of Ebbinghaus (1885) and Schachter (1999, 2001) classed it as one of the seven core “sins” of memory. The decreasing accessibility of memory has been observed over both short (e.g. Peterson & Peterson, 1959) and long (e.g. Wixted & Ebbesen, 1991) intervals, it affects the retention of emotionally salient events (e.g. Talarico & Rubin, 2003) and it has been described as “the culprit in many memory problems” (Schachter, 2001, p. 4).

Explaining the temporal loss of memory is an important priority for cognitive psychologists and there are several theoretical processes that can account for this phenomenon. Common explanations include decay (e.g. Hardt, Nader, & Nadel, 2013), in which a memory gradually deteriorates over time if not used, interference (e.g. Wixted, 2005), in which the memory is distorted or replaced by other information, and context shifts, where changes to the original encoding environment at retrieval may render a memory inaccessible (see Unsworth, Spillers, & Brewer, 2012). Importantly, however, information may disappear from memory at different rates. Fuzzy-trace theory predicts a faster loss of the surface features of experienced objects (verbatim traces) in comparison to less precise gist-based traces (Brainerd & Reyna, 2002). Similarly, the trace transformation hypothesis of systems consolidation proposes that memories are transformed from a detailed and context-specific representation to a schematised version (Winocur & Moscovitch, 2011). The schematised memory lacks specific details but captures the gist of the original event, and memory transformation is thought to occur over time and through experience (Winocur, Moscovitch, & Bontempi, 2010). Both fuzzy-trace theory and the trace transformation hypothesis allow central information or ‘gist’ to survive the passage of time, but finer details are expected to be more rapidly forgotten.

Nonetheless, finely resolved information can persist for some hours, as demonstrated in Brady, Konkle, Alvarez, and Oliva's (2008) study of detailed visual long-term memory (VLTM). Their participants had to view and remember 2,896 pictures of real world stimuli (including 396 repeated images). Each picture showed a single object that was displayed for just 3 s, but the total presentation lasted 5.5 hours. The second phase of the study involved a two-alternative forced-choice (2AFC) recognition test but, crucially, the similarity of the target and foil was manipulated. On novel trials the target and foil were drawn from two distinct categories, permitting the use of low-fidelity representations. Correct responding averaged 92.5%. On exemplar trials, however, the discrimination was more difficult. Here target and foil pictures were drawn from the same category (e.g. two mirrors), but on average participants were correct on 87.6% of the trials. On state trials, the target and foil images both showed the same object, but in a different position or arrangement (e.g. an identical cabinet, but with the doors open or closed). Accurate recognition on state trials required a highly vivid memory, yet participants still performed at a high level (mean correct responses equalled 87.2%).

Brady et al.'s (2008) results show that high-fidelity visual representations can persist for hours, but detailed VLTM has been shown to decline over longer periods of time (e.g. Brady, Konkle, Alvarez, & Oliva, 2013; Hollingworth, 2005). Indeed, Andermane and Bowers (2015) reported equivalent loss of both detailed and gist-based VLTM over one week, contrary to fuzzy-trace and trace transformation theories. Their participants viewed 1,500 images and memory for some of these objects was assessed after a 10-minute or seven-day delay. In a 2AFC recognition task, and following Brady et al. (2008), detailed memory was tested by examining recognition for specific exemplars and object states, and broader gist-level information was assessed on novel trials. All three types of information were forgotten at a similar rate over seven days, with accuracy dropping by approximately 20%. Additionally, Brady et al. (2013) recorded roughly equivalent forgetting rates of state and exemplar information over three days, with recognition accuracy declining by 9.3% on state trials and 12.1% on exemplar trials.

Whilst time-dependent forgetting in VLTM has been clearly documented, there is debate over whether memories are permanently forgotten or just inaccessible at retrieval (see Squire, 2006, for a discussion). In support of the latter possibility, Guerin, Robbins, Gilmore, and Schachter (2012) showed that the loss of finely resolved representations is not inevitable. Guerin et al. had participants study 144 objects that were followed by a recognition test. In one of the recognition conditions, there was a foil related to a previously studied target, but the target itself was not present. Here participants were likely to incorrectly select the foil item that related to the previously studied target, which is known as gist-based false recognition. This mistake occurs when a stimulus similar to one experienced in the past is mistakenly recognised as being identical to the previously encountered item. Yet in another condition, both the target and a similar foil were present and here participants were often able to select the correct object – indeed, there was a 67% drop in gist-based false recognition. This suggests that the details needed to distinguish similar items are still retained and can be reinstated in certain circumstances.

Guerin et al.'s (2012) work also highlighted the important role that retrieval processes play in the subsequent accessibility of memory. Similarly, studies of the testing effect have consistently found that retrieval attempts boost future remembering (Karpicke, 2012; Karpicke & Roediger, 2008; Roediger & Butler, 2011). Whilst the testing effect has typically been reported with verbal materials using recall tasks

(Rowland, 2014), it has been found with abstract visual stimuli (e.g. Kang, 2010) and detailed episodic memories (Sekeres et al., 2016) too. Sekeres et al.'s participants viewed short film clips and had to recall both central and peripheral details following a delay. Peripheral information was lost more quickly than central detail, in line with fuzzy-trace theory and the trace transformation hypothesis, but retrieving the clips after encoding preserved *both* peripheral and central details across a seven-day interval.

The findings reviewed above suggest that detailed memories may be capable of surviving the passage of time if retrieval practice is provided, and the present study aimed to provide a strong test of this idea by investigating the retention of exemplar and state information over a week. Memory for these properties does not rely on gist (Brady et al., 2008) and is susceptible to time-dependent forgetting (Andermane & Bowers, 2015), hence there would be important theoretical implications if an initial retrieval attempt (without feedback) preserved such detailed memory.

Exploring memory for state and exemplar information also allowed the nature of visual representations to be investigated. It is possible that items in VLTM are held as bound units or individual objects (see Brady et al., 2013, for an overview of this account). If so, all components of a visual memory should collapse at the same rate, and as already discussed some evidence is consistent with this view (e.g. Andermane & Bowers, 2015). Other theorists have argued against an "all-or-nothing" approach to forgetting and proposed that the various components of a memory may be lost at different rates (e.g. Sekeres et al., 2016). In VLTM, this view rejects the notion of bound representations and allows particular elements of a representation (e.g. colour or orientation) to be lost independently. Brady et al. (2013) tested this account by asking participants to remember 120 object images and then complete a surprise recognition test either immediately after the object viewing (short delay) or following a three-day interval. The recognition test involved a four-alternative forced-choice procedure, where two exemplars were shown in two states. Participants had to select the specific image they had seen before, but Brady et al. were interested in the dependence between state and exemplar information (i.e. the extent to which correct identification of the exemplar would also lead to correct identification of its state).

Data were assessed using a dependence score that corrected for guessing, with 100% denoting complete dependence (e.g. if exemplar detail was remembered, state detail would also always be remembered) and 0% denoting complete independence (e.g. remembering exemplar detail would confer no benefit on remembering state information). Following an immediate delay, dependence scores were 46.6% for state dependence on exemplar information, and 27.4% for exemplar dependence on state information. Three days later, these two dependence scores had decreased to 13.4% and 7.6%, respectively, indicating increased independence of the two properties over time. Brady et al. argued that state and exemplar information are forgotten independently as time passes. For instance, participants may have remembered that a glass of orange was half empty (state information), but did not recall the specific glass (exemplar information), or they may have remembered a specific breakfast cereal container (exemplar information), but forgot that the container was open, rather than closed (state information). Aside from challenging the notion of bound visual representations, Brady et al.'s results highlight that forgetting can affect specific properties of a memory.

Existing explanations of time-based forgetting, such as fuzzy-trace theory and the trace transformation hypothesis, or models relying on decay, interference and context shifts, do not necessarily anticipate specific components of detailed representations (e.g. the precise exemplar and its state) to be lost in different ways. The evidence concerning this possibility is mixed and relatively little work has addressed forgetting of state- and exemplar- information over prolonged delays. Consequently, the present experiment aimed to both chart the loss of state and exemplar information over seven days and assess whether these two types of VLTM responded differently to an initial retrieval attempt.

In Experiment 1, 300 images were presented to participants and followed by an immediate 2AFC recognition test that assessed exemplar and state memory. The second recognition test occurred one week later. Importantly, some participants were tested on the same stimuli during both phases, providing retrieval practice, whereas other participants were tested on different sets of stimuli (this served as a control or baseline condition). Most theories of time-based forgetting predict a loss in accuracy over one week, particularly given the number of stimuli involved and the fine details needed to perform the recognition task. However, inspired by the testing effect and other relevant findings (e.g. Sekeres et al., 2016), a preliminary retrieval attempt was expected to preserve the detailed information needed to distinguish the target from the foil, making this representation more accessible in the future. Whether retrieval practice would differentially affect exemplar and state memory was unclear, but Brady et al.’s (2013) notion of independent forgetting would be supported if state and exemplar information did react differently.

EXPERIMENT 1

Method

Participants

One hundred and sixty (134 females and 26 males) first-year undergraduate psychology students from the University of Wolverhampton completed both phases of the experiment. Participants were aged between 18 and 58 (*M* = 21.49, *SD* = 6.71) and did the experiment as part of an introductory research methods session. By recruiting participants through a psychology course, it was possible to obtain a much larger sample than has been used in studies of a similar nature (e.g. Andermane & Bowers, 2015; Brady et al., 2008), hence reducing concerns about low statistical power. An additional 44 participants completed phase 1 only and were removed from the analysis. All individuals gave written informed consent before beginning the experiment and the study protocol was approved by a Faculty ethics committee.

Materials

Three hundred images were used during the picture presentation, including 285 pictures of real world objects and animals (64 targets and 221 fillers). The fillers comprised images randomly selected from databases provided by Brady et al. (2008, 2013, see http://konklab.fas.harvard.edu/#). Effort was taken to ensure that real world objects were sufficiently different from one another and images were categorically distinct. Additionally, 15 artificial, computer-generated images were randomly chosen from McKeown, Holt, Delvenne, Smith, and Griffiths' (2014) stimulus set.

For the recognition test, each target was presented alongside a foil (see Figure 1). There were 32 exemplar pairs and 32 state pairs, randomly selected from Brady et al.'s (2008) database. On exemplar trials, the foil was drawn from the same category as the target picture, but was a different item (e.g. two TV remotes). On state trials, the foil was the same item as the target, but presented in a different position or arrangement (e.g. a clipboard with and without a pen). One object was labelled "M" and the other was labelled "Z", with the target and foil being presented in each position an equal number of times. Lastly, 16 of the state and exemplar recognition trials were randomly allocated into one stimulus set (Set A), with the remaining trials forming another block (Set B).

"Figure 1 about here"

Design and Procedure

The study employed a 2 (phase: 1 vs. 2) x 2 (trial type: state vs. exemplar) x 2 (condition: retrieval practice vs. baseline) mixed design. Both phases of the experiment were carried out in a large classroom under exam conditions with four separate groups. Responses were collected on paper. Prior to the initial picture presentation, participants were seated in rows facing a large screen and were told to try and remember as many images as possible. They were not told about the need to retain state- or exemplar-level detail but they were asked to tally the number of computer-generated items that were displayed. This was intended to help participants maintain concentration and attend to the images. During the presentation, each image was shown in the centre of the screen for 3 s and followed by a 1 s blank interval.

The first recognition test occurred immediately after the presentation. The arrangements of the recognition task and the required responses were explained to participants. They were told that the target and foil could be very similar, but again the state- and exemplar-level difference was not explicitly outlined. Trials were shown for 5 s and succeeded by a 3 s unfilled interval. Participants had to select the image they had seen before. Exactly seven days later, participants returned to the laboratory and completed the second phase of the study, which consisted of the recognition test only. Importantly, some participants repeated the phase 1 recognition test (retrieval practice condition), whereas other participants were given a new set of recognition trials (baseline condition). The combination of stimulus sets was balanced across the four groups (i.e. the retrieval practice condition experienced either Set A twice, or Set B twice, whereas the baseline condition experienced Set A followed by Set B, or vice versa).

Results

Participants were excluded if six or more responses were missing during either recognition test. This affected two individuals, so the final sample included 158 participants (retrieval practice: N = 78; baseline: N = 80). Other missing responses were treated as incorrect, but these were rare and only affected four individuals.

The mean proportion of correct responses is shown in Figure 2 for all conditions. A three-way mixed ANOVA was used to assess these data, with trial type (state vs. exemplar) and phase (1 vs. 2) as within-subject factors, and condition (retrieval practice vs. baseline) as the between-groups factor[1]. Significant interactions were followed up with simple effects analysis and independent-groups *t*-tests, all corrected using the Holm-Šidàk technique.

All main effects were significant. Performance at phase 1 (M = .79) exceeded that at phase 2 (M = .68), $F(1, 156) = 173.20$, MSE = .01, $p < .001$, $\eta_p^2 = .53$, participants were slightly more accurate on state (M = .75) than exemplar (M = .72) trials, $F(1, 156) = 10.70$, $MSE = .01$, $p = .001$, $\eta_p^2 = .06$, and performance was better in the retrieval practice ($M = .75$), than baseline ($M = .71$) condition, $F(1, 156) = 6.20$, $MSE = .04$, $p = .014$, $\eta_p^2 = .04$.

The phase by trial type interaction was non-significant, $F(1, 156) = .25$, $MSE = .01$, $p = .615$, $\eta_p^2 < .01$, but all other interactions were reliable, including the theoretically relevant interaction between phase and condition. During phase 1, there was no significant difference between the retrieval practice and baseline conditions (retrieval practice: $M = .78$; baseline: $M = .80$, $t[156] = -.78$, $p = .436$, $d = -.17$), whereas during phase 2 performance was higher in the retrieval practice group (retrieval practice: $M = .72$; baseline: $M = .63$, $t[156] = 5.47$, $p < .001$, $d = .82$).

The other significant interactions were subsumed by a higher order three-way interaction, $F(1, 156) = 15.91$, $MSE = .01$, $p < .001$, $\eta_p^2 = .09$, which was explored with separate 2 (phase: 1 vs. 2) x 2 (condition: retrieval practice vs. baseline) mixed ANOVAs for state and exemplar trials. The analysis for state information showed a significant main effect of phase only, $F(1, 156) = 89.13$, $MSE = .01$, $p < .001$, $\eta_p^2 = .36$, with accuracy declining from 80.2% in phase 1 to 69.3% in phase 2, on average. The analysis on exemplar trials also showed a significant effect of phase, $F(1, 156) = 95.42$, $MSE = .01$, $p < .001$, $\eta_p^2 = .38$, with a decline from phase 1 ($M = .78$) to 2 ($M = .66$), and a significant effect of condition, $F(1, 156) = 11.48$, $MSE = .03$, $p = .002$, $\eta_p^2 = .07$. Performance was better in the retrieval practice ($M = .75$) than baseline ($M = .69$) condition, and the interaction was also significant, $F(1, 156) = 54.41$, $MSE = .01$, $p < .001$, $\eta_p^2 = .26$. This was followed up with independent *t*-tests, which as expected showed no difference between conditions at phase 1, $t(156) = -1.01$, $p = .314$, $d = -.21$. However, during phase 2 participants were significantly more accurate if they had experienced retrieval practice (*M* = .74), rather than being tested on a new set of exemplar trials, (*M* = .58), $t(156) = 6.97$, $p < .001$, $d = 1.14$.

“Figure 2 about here”

The final analysis more directly explored the forgetting rate. To calculate forgetting, scores at phase 2 were subtracted from those at phase 1 for state and exemplar trials. A 2 (trial type: state vs. exemplar) x 2 (condition: retrieval practice vs. baseline) mixed ANOVA showed that accuracy declined more in the baseline (*M* = .17) than retrieval practice (*M* = .06) condition, $F(1, 156) = 43.02$, $MSE = .02$, $p < .001$, $\eta_p^2 = .22$, but there was no main effect of trial type, $F(1, 156) = .25$, $MSE = .02$, $p > .250$, $\eta_p^2 = .00$. However, there was a significant interaction, $F(1, 156) = 15.91$, $MSE = .02$, $p < .001$, $\eta_p^2 = .09$. For the retrieval practice condition, state trials (*M* = .08) showed a greater decline in accuracy over the delay than exemplar trials (*M* = .03), $t(77) = 2.61$, $p = .011$, $d = .31$. In the baseline condition the pattern was reversed, with exemplar information being forgotten more rapidly (*M* = .20) than state detail (*M* = .13), $t(79) = -3.03$, $p = .006$, $d = -.33$.

Discussion

An initial retrieval attempt protected detailed VLTM from time-based forgetting over one week, but this effect was mainly driven by memory for object exemplars. The loss of exemplar detail over the delay decreased from approximately 20% in the baseline condition (replicating Andermane & Bowers, 2015) to just 3% in the retrieval practice condition. State information was forgotten at a similar rate regardless of retrieval practice, suggesting an important difference in the way state and exemplar information is retained over time. However, the latter finding should be treated with some caution. Whilst state information was susceptible to time-dependent forgetting in both the retrieval practice and baseline groups, there was a trend towards better performance following retrieval practice at phase 2.

In addition, performance in the retrieval practice condition may have been influenced by the placement of the target and foil, as they were presented in the same location in both recognition tests. The survival of exemplar information over the delay may therefore have been influenced by the target placement and associated spatial memories, at least to some degree. Another influential variable concerned the foil. During phase 1 participants in the retrieval practice group may have formed a detailed memory of the wrong object. In the subsequent recognition test, participants may have struggled to differentiate the target from the foil, at least on state trials. Experiment 2 was designed to address these issues.

EXPERIMENT 2

Experiment 2 matched some of the arrangements of Experiment 1, but here all participants were given retrieval practice. The intention was to replicate some of the findings emerging from the first study, whilst addressing some of the problems and alternative explanations of the retrieval practice effect. Firstly, to control for responding based on object placement rather than identity, half of the target and foils switched position during the phase 2 recognition test. Secondly, to assess the impact of foil interference, three different foil types were developed.

On “old” trials, identical target and foils were provided during both phases, replicating Experiment 1. Following the results of that experiment, accuracy on old state trials was expected to decline over the interval, whereas exemplar information was predicted to be relatively robust and resistant to temporal forgetting. Yet these old trials may be particularly susceptible to foil interference, which may lessen the advantage of retrieval practice (and particularly when the placement of the target varies across phases). To examine the impact of foil interference, “new” recognition trials were created. Here, two different foils were used during phase 1 and 2, and if performance is affected by interference from memory for an old foil, accuracy on new trials should show a reduced decline over the interval.

The final foil type created a “switch” in the information being tested. On exemplar switch trials, a phase 1 exemplar trial would be converted to a state trial during phase 2, and vice versa on state switch trials. At one level, switch trials acted in a similar way to new trials and assessed whether forgetting over time could be reduced when foil interference was eliminated. At a deeper theoretical level, switch trials aimed to determine whether retrieval practice for one type of information (state or exemplar) could preserve other information over a prolonged delay. For instance, if exemplar information can survive following retrieval practice, converting a state trial into an exemplar trial should allow accuracy to persist at a high level. Yet if state information is inevitably lost over time, converting an exemplar trial into a state trial should offer no protection against time-dependent forgetting. Alternatively, it is possible that the benefits of retrieving exemplar information during phase 1 can generalise and preserve some information about the object state, allowing state information to be retained and reinstated in future.

Method

Participants

One hundred and thirty-one first-year undergraduate psychology students (115 females and 16 males) from the University of Wolverhampton completed both phases of the experiment and submitted valid responses. Participants were aged between 17 and 53 (*M* = 23.40, *SD* = 7.47) and completed the experiment during an introductory research methods session. None of the participants had taken part in Experiment 1. Thirty-three participants completed phase 1 only and were removed from the analysis. Two additional participants withdrew their data and another three did not provide the code that allowed responses in the two phases to be linked. All individuals gave written informed consent before beginning the experiment and the study protocol was approved by a Faculty ethics committee.

Materials

A total of 300 categorically distinct and unique images (72 targets, 212 fillers and 16 artificial stimuli) were used in the picture presentation, with fillers again being selected from Brady et al. (2008, 2013) and artificial stimuli being selected from McKeown et al. (2014). For the recognition task, each condition included 12 trials. Twelve state and 12 exemplar trials were randomly selected from those used in Experiment 1 to make the “old” trials. To create the “new” recognition trials for exemplar stimuli, pictures were randomly selected from Konkle, Brady, Alvarez, and Oliva’s (2010) database. Twelve categories were selected for use here, with three images (one target and two foils) being chosen for the recognition test. For the “new” state trials, the Bank of Standardized Stimuli (BOSS) was used (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010). The database was searched for images that had at least three photos of the same object, but in different arrangements or positions. This created nine suitable trials. The remaining three trials were established by selecting images of the same object in three states using online searches. Lastly, Brady et al.’s (2013) database helped develop the “switch” trials. This database contains four items (two exemplars in two states) for 100 objects. Stimuli were randomly selected to form the switch trials used here. See Figure 3 for examples. Other arrangements matched Experiment 1, except the target and foil swapped position on half of the recognition trials in phase 2.

“Figure 3 about here”

Design and Procedure

The study employed a 2 (phase: 1 vs. 2) x 2 (trial type: state vs. exemplar) x 3 (foil type: old vs. new vs. switch) repeated measures design. The procedure matched that used in Experiment 1, but given the fully within design there was no need to place recognition trials into different sets.

Results

As in Experiment 1, missing responses were treated as incorrect but only three participants omitted a response and none had more than one missing answer. Mean proportion of correct responses is shown in Figure 4 for all conditions. The data were explored using ANOVA, with Šidàk post-hoc tests being used to follow up the main effect of recognition trial type. Violations to sphericity were corrected using the Greenhouse-Geisser adjustment and specific comparisons needed to interpret interactions were achieved using Holm-Šidàk corrected *t*-tests.

A two-way repeated measures ANOVA was performed, with phase and trial type (exemplar old vs. exemplar new vs. exemplar switch vs. state old vs. state new vs. state switch) as factors. Exemplar and state trials were collapsed into a single variable in this analysis because of the switch condition (as this condition involved both exemplar and state information, it was difficult to separate state/exemplar variables in the model). However, subsequent analyses allowed more specific comparisons to be made.

This ANOVA found both main effects and the interaction to be significant. Accuracy modestly declined from 82.5% in phase 1 to 80.1% in phase 2, $F(1, 130) = 17.44$, $MSE = .01$, $p < .001$, $\eta_p^2 = .12$, and differences were found between the six trial types, $F(5, 650) = 83.76$, $MSE = .01$, $p < .001$, $\eta_p^2 = .39$. Performance on exemplar new ($M = .90$) and exemplar switch ($M = .89$) trials was notably higher than all other trial types ($p < .001$), whereas poorest performance was recorded on state trials with a new foil ($M = .74$). Indeed, accuracy on these trials was reliably worse than all other trial types ($p < .001$), with the exception of exemplar trials with an old foil ($M = .76$, $p = .840$). Performance on these exemplar old trials was significantly lower than state switch ($p = .042$) and marginally lower than state old trials ($p = .062$), whereas the latter two conditions did not differ.

The significant phase x trial type interaction, $F(5, 650) = 20.30$, $MSE = .01$, $p < .001$, $\eta_p^2 = .14$, suggested differences in the forgetting rate according to the type of information being tested (state or exemplar) and the nature of the foil. To better understand the interaction, 2 (phase: 1 vs. 2) x 2 (trial type: state vs. exemplar) ANOVAs were performed separately for the three foil types.

For new foils, there was no effect of phase, $F(1, 130) = .83$, $MSE = .01$, $p = .363$, $\eta_p^2 = .01$, but accuracy on exemplar trials ($M = .90$) was noticeably higher than state trials ($M = .74$), $F(1, 130) = 252.70$, $MSE = .01$, $p < .001$, $\eta_p^2 = .66$. The interaction was also significant, $F(1, 130) = 12.12$, $MSE = .01$, $p = .001$, $\eta_p^2 = .09$, with significant time-dependent forgetting on state trials, $t(130) = 2.89$, $p = .008$, $d = .26$. Specifically, there was a modest decline of around 4% over the one-week delay, but on exemplar trials the opposite pattern was seen, $t(130) = -2.10$, $p = .038$, $d = -.15$. Accuracy modestly increased from 88.78% at phase 1 to 91.16% at phase 2.

For switch foils, there was a significant effect of phase, $F(1, 130) = 4.98$, $MSE = .01$, $p = .027$, $\eta_p^2 = .04$, but this was due to improved performance at phase 2 ($M = .83$), in comparison to phase 1 ($M = .85$). The effect of trial type was also significant, $F(1, 130) = 107.58$, $MSE = .01$, $p < .001$, $\eta_p^2 = .45$, but this is less informative as switch trials compare different types of information over the delay (e.g. a phase 1 exemplar is compared against a phase 2 state, or vice versa). The interaction was non-significant, $F(1, 130) = .79$, $MSE = .01$, $p = .374$, $\eta_p^2 = .01$, showing that both state and exemplar switch trials had improved performance over the delay (when state switched to exemplar, accuracy increased by 1.28%, and when exemplar switched to state, accuracy improved by 2.6%). Whilst these improvements are modest, it does reflect better recognition performance after a seven-day delay.

For old foils, there was clearer evidence of time-dependent forgetting, with performance dropping by around 8%, on average (phase 1: $M = .82$; phase 2: $M = .74$), $F(1, 130) = 94.26$, $MSE = .01$, $p < .001$, $\eta_p^2 = .42$. Accuracy on state trials ($M =$

.79) also exceeded that on exemplar trials (*M* = .76), $F(1, 130) = 8.48$, $MSE = .02$, $p = .004$, $\eta_p^2 = .06$. However, the interaction was non-significant, $F(1, 130) = .03$, $MSE = .01$, $p = .864$, $\eta_p^2 = .00$, suggesting similar loss of both state and exemplar detail. Thus, unlike Experiment 1, there was time-dependent forgetting on exemplar old trials. This may have been influenced by improved control over responding based on spatial location (specifically, the target and foil could switch position in the two recognition tests, unlike Experiment 1).

To investigate any effect of the target and foil position on exemplar old trials, a 2 (phase: 1 vs. 2) x 2 (position of target during phase 1 and 2: same vs. different) repeated measures ANOVA was conducted. As before, the effect of phase was significant, $F(1, 130) = 49.62$, $MSE = .02$, $p < .001$, $\eta_p^2 = .28$, with performance declining over the delay, but position was also significant, $F(1, 130) = 35.03$, $MSE = .03$, $p < .001$, $\eta_p^2 = .21$. There was higher accuracy on trials where the target and foil remained in the same position (*M* = .81), rather than a different position (*M* = .72). Of most relevance, the interaction was significant, $F(1, 130) = 4.71$, $MSE = .02$, $p = .032$, $\eta_p^2 = .04$, and is shown in Table 1. Accuracy was higher when the target and foil was in the same position, but this applied to both recognition phases, suggesting that the trials where the items switched position may have been more difficult overall. However, the forgetting rate was also greater on trials where the target and foil switched position (approximately 11%, compared to 5.5% on same trials, on average).

"Table 1 about here"

Discussion

The results of Experiment 2 were broadly compatible with those in Experiment 1. Exemplar-level memory appeared more robust over the delay than state-level memory but, unlike Experiment 1, old exemplar trials (where the same target and foil were used at both phases) showed a significant decline over the interval. The major difference between this trial type in the two experiments was the control over object placement – in Experiment 1, the target and foil were placed in the same position during both phases, but this was not always the case in Experiment 2. This suggests that the benefits of retrieval practice on exemplar trials in Experiment 1 may have been influenced by memory for the old target position, rather than the precise object. When this possibility was controlled in Experiment 2, there was modest but significant forgetting of exemplar detail (and this was greater on trials where the target and foil switched position). Yet where foil interference was controlled, exemplar memory did survive over the interval. Providing a new exemplar foil led to a slight *improvement* in accuracy after a week had passed, and a similar absence of forgetting was seen on the state switch trials. Practicing the retrieval of state information during phase 1 also seemed to preserve memory for exemplar detail one week later.

Conversely, accuracy on state trials decreased over time. This occurred regardless of whether the original foil or a new foil was presented during phase 2 (showing that any forgetting was not simply the result of interference from the original foil), although forgetting on state trials with a new foil seemed to be relatively minor. Furthermore, there was an absence of forgetting in the exemplar switch condition. If participants were tested on the object exemplar during phase 1, switching to a state trial for that same object at phase 2 preserved performance (indeed, there was a small but significant recovery in accuracy during the delayed recognition test). Thus, even very detailed state memory can survive over one week if the initial retrieval attempt tests exemplar memory.

EXPERIMENT 3

The first two experiments suggested that retrieval practice may preserve detailed VLTM over one week, although the precise effect is dependent on the information being retained and the type of foil that is used. In Experiment 1, retrieval practice slowed the forgetting of exemplar (but not state) information, in comparison to a baseline condition. In Experiment 2, practicing the retrieval of some stimuli during phase 1 prevented time-based forgetting (though this was not seen on “old” or “state new” trials). The overall results show that some forms of detailed VLTM can survive over time, but it is unclear whether retrieval practice is entirely responsible for this effect.

Any apparent benefit of retrieval practice may actually be due to mere exposure. In the retrieval practice condition, participants experienced the target stimuli twice – once during the initial presentation and again during the phase 1 recognition test. It could be that the simple repetition of the target strengthened the memory and allowed it to persist over time. In standard testing effect studies, mere exposure is controlled by comparing the retrieval practice group against a restudy group, where information is re-encoded but not deliberately retrieved (Rowland, 2014). Evidence for a testing effect is only accepted where performance on a final test is better in the retrieval practice than the restudy condition. This finding is typical in the testing effect literature (Rickard & Pan, 2017), suggesting that the act of retrieval practice is crucial, but without an appropriate control it is difficult to draw firm conclusions for the present study.

So, in Experiment 3 a new “restudy” condition was added to the design. After the initial presentation, participants in this group saw the 64 images that would be tested during phase 2 for a second time. However, they were tested on non-repeated images during the phase 1 test, preventing any retrieval practice of the critical stimuli. The restudy group was then compared against retrieval practice and baseline conditions. It was anticipated that retrieval practice would lead to higher accuracy during phase 2 in comparison to both the restudy and baseline conditions, although repeated encoding could produce some modest facilitation, in comparison to baseline.

Experiment 3 also aimed to assess the impact of trial type and foil type on recognition performance, given some mixed results. Whilst Experiment 1 did not uncover any loss of exemplar information when the same foil was used in both phases, Experiment 2 did. As noted above, one possible explanation for this discrepancy was better control over object placement in the second experiment, where the target and foil sometimes switched position across phases following retrieval practice. A specific analysis on exemplar old trials incorporating position information supported this interpretation, where there was greater forgetting on trials where the target and foil swapped position. In addition, some of the exemplar trials in Experiment 1 also varied object state, whereas the present experiment prevented this. Experiment 3 therefore allowed another opportunity to assess the robustness of exemplar memory following retrieval practice, in relation to both baseline and restudy conditions. The nature of time-dependent forgetting in state memory was also assessed, as some of the results from Experiment 2 indicated that state information may survive the passage of time in some circumstances.

Finally, Experiment 3 reintroduced the switch trials. In Experiment 2, this condition yielded some surprising results, as it highlighted modest facilitation over the delay when a state trial was converted to an exemplar trial, and vice versa. The replicability of this finding was assessed in the present experiment, and the switch trials also served as a tool for monitoring foil interference, which may play an important role. Due to the nature of switch trials, “true” switches could only occur in the retrieval practice group (in order for a noticeable switch to happen, participants must be tested on different types of information for the same target). In the restudy and baseline groups, switch trials involved the comparison of exemplar and state, or state and exemplar information, but for different stimuli. As such, the restudy and baseline groups acted as a comparison against which the impact of switch trials in the retrieval practice condition could be measured. Importantly, the specific information tested on all trial types during phase 2 was identical in the three conditions – for example, a phase 2 state switch trial presented the same exemplar trial for retrieval practice, restudy and baseline groups. Recognition accuracy during phase 2 was therefore of particular interest in Experiment 3.

Method

Participants

In total, 153 first-year undergraduate psychology students (137 females, 12 males and four unreported sex) from the University of Wolverhampton completed both phases of the experiment. Participants were aged between 18 and 59 (*M* = 23.91, *SD* = 8.23) and undertook the experiment during an introductory research methods session. None of the participants had taken part in the previous experiments. In addition, twenty-two participants completed phase 1 only, three participants withdrew their data, four did not provide the code that allowed responses to be linked and two self-reported visual impairments. All were removed from the analysis. Every participant gave written informed consent before beginning the experiment and the study protocol was approved by a Faculty ethics committee.

Materials

As in Experiments 1 and 2, 300 categorically distinct and unique images (128 targets, 156 fillers and 16 artificial stimuli) were used in the picture presentation and selected from Brady et al. (2008, 2013) and McKeown et al. (2014). The recognition task contained 64 trials during each phase (16 for each state/exemplar and old/switch trial type), and the target and foil swapped position on half of the recognition trials in the phase 2 retrieval practice condition.

To create better control over the state/exemplar manipulation, stimuli from Brady et al.'s (2013) database were used for the retrieval practice group (both phases) and the phase 2 test for restudy and baseline conditions. Trials in the phase 2 recognition test were therefore identical in all three groups. Ninety-six of the 100 objects from Brady et al.'s database were selected for use here (the remaining four objects were removed as they may have been unfamiliar to a UK sample). Importantly, exemplar recognition pairs in this database show objects in approximately the same state, and it was a suitable database for creating switch trial stimuli, as it contains state and exemplar representatives for each item. During phase 1, the restudy and baseline recognition test included 32 stimuli from Brady et al.

(2013) and, due to an absence of sufficient stimuli, 32 items from Brady et al. (2008). However, these latter stimuli were selected to ensure that exemplar recognition pairs were in approximately the same state. The restudy and baseline groups were tested on different stimuli in phases 1 and 2, matching the baseline condition of Experiment 1. Nonetheless, care was taken to ensure any comparisons between phases remained meaningful. For instance, old trials tested the same type of memory (exemplar or state) during both phases, but for different stimuli. Switch trials compared memory on phase 1 state trials against memory for different phase 2 exemplar trials, and vice versa.

Design and Procedure

The study manipulated four variables in a 2 (phase: 1 vs. 2) x 2 (trial type: state vs. exemplar) x 2 (foil type: old vs. switch) x 3 (condition: retrieval practice vs. restudy vs. baseline) mixed design. The condition variable was a between-subjects factor and all other variables were manipulated within groups.

Both phases of the experiment took place in the same classroom under exam conditions, with images projected onto a large screen. Six separate groups were tested, and these were randomly allocated to either the retrieval practice, restudy or baseline conditions. To reduce picture-specific effects, a counterbalancing arrangement was employed. Within each condition, the stimuli that comprised exemplar old trials for one group acted as exemplar switch trials for the other group, and vice versa. The same arrangement was adopted for state stimuli. During phase 1, the retrieval practice condition was tested on a different set of items to the restudy and baseline groups, but during phase 2 all conditions completed the recognition test for the same stimuli.

The procedural arrangements for the retrieval practice and baseline conditions matched Experiment 1, but the restudy group had an added encoding opportunity. Following the presentation of the 300 images, the 64 targets that would be tested at phase 2 were shown for a second time. They were presented at the standard rate (3 s and followed by a 1 s gap), and participants were informed in advance that stimuli would be shown either once or twice, so some repetition was expected. They then completed the phase 1 recognition task on a different set of stimuli (i.e. repeated images were only tested at phase 2).

The restudy condition was designed to control for any effect of re-exposure, but the retrieval practice group were shown the target for 5 s during the recognition test. Whilst this created a slightly longer exposure time in the retrieval practice condition, during this 5 s participants had to consider the foil item and decide upon the image they had seen before. The 3-s presentation in the restudy condition therefore seemed like a reasonable approximation to exposure to the target stimuli in the retrieval practice group.

Results

Following Experiments 1 and 2, participants were removed if they had six or more missing responses, or where responses had become confused and out of sequence, and thus difficult to score. This affected eight individuals in total. The final sample comprised 145 participants, with 46, 44 and 55 individuals in the retrieval practice, restudy and baseline conditions, respectively.

Data were assessed using ANOVAs, with simple effects and Holm-Šidàk corrected *t*-tests used to follow up significant interactions. However, the switch from exemplar to state, or state to exemplar, only occurred in a meaningful way in the retrieval practice condition. In the restudy and baseline conditions, participants had no experience of any switch, as this was the first time they were tested on those specific items. Similarly, as in Experiment 2, the presence of switch trials made it difficult to differentiate state from exemplar in the ANOVA, as switch trials include both types of information. To address these difficulties, the trial type (state/exemplar) and foil type (old/switch) variables were treated as one factor with four levels, but subsequent analyses allowed more specific comparisons to be made.

A 2 (phase: 1 vs. 2) x 3 (condition: retrieval practice vs. restudy vs. baseline) x 4 (trial: state old vs. state switch vs. exemplar old vs. exemplar switch) mixed ANOVA was then performed. Aside from a null effect of condition, $F(2, 142) = 1.11$, $MSE = .07$, $p = .332$, $\eta_p^2 = .02$, all other main effects and interactions were significant. Mean accuracy dropped from 77.2% to 67.1% over the delay, $F(1, 142) = 159.75$, $MSE = .02$, $p < .001$, $\eta_p^2 = .53$, and differences were found between the four trial types, $F(3, 426) = 14.56$, $MSE = .01$, $p < .001$, $\eta_p^2 = .09$. Šidàk post-hoc tests showed performance on state old trials ($M = .69$) to be worse than all other trial types (all *p*s < .008). Exemplar old trial accuracy ($M = .72$) was also lower than both switch trials (*p*s < .009), although there was no significant difference between exemplar switch and state switch trials ($M = .74$, $p = .720$).

The theoretically relevant phase x condition interaction was significant, $F(2, 142) = 30.84$, $MSE = .02$, $p < .001$, $\eta_p^2 = .30$, and is shown in Table 2. No reliable differences were found among the three conditions at phase 1, $F(2, 142) = 1.89$, $MSE = .03$, $p = .154$, but there was an effect at phase 2, $F(2, 142) = 16.65$, $p < .001$.

Accuracy in the retrieval practice condition exceeded both the baseline, $t(99) = 5.76$, $p < .001$, $d = 1.18$, and restudy, $t(88) = 2.04$, $p = .045$, $d = .44$, conditions. The restudy group also outperformed the baseline condition, $t(97) = 3.49$, $p = .002$, $d = .71$.

"Table 2 about here"

Both the phase x trial type interaction, $F(3, 426) = 31.28$, $MSE = .01$, $p < .001$, $\eta_p^2 = .18$, and trial by condition interaction, $F(6, 426) = 2.42$, $MSE = .01$, $p = .026$, $\eta_p^2 = .03$, were significant. The former interaction suggested that the forgetting rate depended upon the type of information being tested, whereas the latter interaction highlighted differences in trial performance according to condition. However, both of these interactions were subsumed by a higher order three-way interaction, $F(3, 426) = 7.64$, $MSE = .01$, $p < .001$, $\eta_p^2 = .10$, and given that genuine switch trials only occurred in the retrieval practice condition, only the three-way interaction was explored further. This higher order interaction was broken down to 1) assess time-dependent forgetting and 2) explore differences between the three conditions at phase 2, with this latter analysis being of most interest.

To examine time-dependent forgetting, a 2 (phase: 1 vs. 2) x 4 (trial: state old vs. state switch vs. exemplar old vs. exemplar switch) repeated measures ANOVA was conducted separately for each condition. For all analyses, both main effects and the interaction were significant[2], so paired *t*-tests were used to compare phase 1 and 2 performance for each trial type, separately in the three conditions (see Table 3). In the retrieval practice condition, all trial types declined over the delay, except for the state switch condition where there was an improvement in recognition accuracy during phase 2. In the restudy and baseline conditions, old trials compared memory for the same type of information across a delay (e.g. phase 1 state vs. phase 2 state), whereas switch trials compared independent exemplar and state trials (e.g. phase 1 exemplar vs. phase 2 state). In both conditions, the specific images tested during phase 1 were unrelated to those tested during phase 2, unlike the retrieval practice condition. Robust time-dependent forgetting occurred for all trial types, with the only exception being exemplar old trials in the restudy condition.

"Table 3 about here"

To explore differences between the three conditions at phase 2, where all participants were responding to the same stimuli, a 2 (trial type: state vs. exemplar) x 2 (foil type: old vs. switch) x 3 (condition: retrieval practice vs. restudy vs. baseline) mixed ANOVA was used. The effect of old and switch trials in the retrieval practice condition were assessed in relation to restudy and baseline groups, which acted as comparisons.

Higher accuracy was found on exemplar (*M* = .70) than state (*M* = .64) trials, $F(1, 142) = 59.55$, $MSE = .01$, $p < .001$, $\eta_p^2 = .30$, and on switch (*M* = .69) than old (*M* = .65) trials, $F(1, 142) = 25.97$, $MSE = .01$, $p < .001$, $\eta_p^2 = .16$. The interaction between trial type and condition, $F(2, 142) = 1.45$, $MSE = .01$, $p = .239$, $\eta_p^2 = .02$, and between trial type and foil type, $F(1, 142) = .03$, $MSE = .01$, $p = .855$, $\eta_p^2 = .00$, were both non-significant. However, there was an overall effect of condition, $F(2, 142) = 16.64$, $MSE = .03$, $p < .001$, $\eta_p^2 = .19$, with both the retrieval practice (*M* = .72, $p < .001$) and restudy (*M* = .68, $p = .002$) conditions significantly outperforming the baseline group (*M* = .62), but not reliably differing from each other ($p = .107$). Additional interactions were found between foil type and condition, $F(2, 142) = 4.53$, $MSE = .01$, $p = .012$, $\eta_p^2 = .06$, and amongst all three variables, $F(2, 142) = 12.02$, *MSE* = .01, $p < .001$, $\eta_p^2 = .15$. This higher order three-way interaction was followed up with a series of one-way ANOVAs comparing the three groups on each trial type. All tests were significant (all *p*s < .008), but differences between the groups were found through *t*-tests.

On state old trials, higher accuracy was found in the retrieval practice (*M* = .67) than the restudy (*M* = .60, *t*[88] = 2.28, *p* = .049, *d* = .43) and baseline (*M* = .57, *t*[99] = 4.00, *p* < .001, *d* = .74) conditions. The latter two groups did not differ (*t*[97] = 1.36, *p* = .176, *d* = .21). On exemplar switch trials, also assessing state memory, the retrieval practice group (*M* = .70) performed better than baseline (*M* = .62, *t*[99] = 3.13, *p* = .006, *d* = .67), but not restudy (*M* = .67, *t*[88] = .90, *p* = .373, *d* = .25). There was a marginal difference when comparing the latter two groups, but this was not significant, *t*(97) = 2.17, *p* = .065, *d* = .42).

There was no beneficial effect of retrieval practice on exemplar old trials, with the only difference occurring between restudy (*M* = .73) and baseline (*M* = .64), *t*(97) = 3.41, *p* = .003, *d* = .72. Yet on state switch trials, assessing exemplar memory, the retrieval practice group (*M* = .83) outperformed both the restudy (*M* = .71, *t*[88] = 4.51, *p* < .001, *d* = 1.00) and baseline (*M* = .64, *t*[99] = 8.43, *p* < .001, *d* = 1.73) conditions. Higher accuracy was also seen in the restudy than baseline groups, *t*(97) = 2.76, *p* = .007, *d* = .54.

Finally, to assess any influence of the target and foil spatial location, a separate 2 (phase: 1 vs. 2) x 2 (position of target during phase 1 and 2: same vs. different) x 4 (trial type: exemplar old vs. exemplar switch vs. state old vs. state switch) repeated measures ANOVA was carried out on the retrieval practice condition. This condition was chosen as it was the only group in which the position of the target and foil could be reliably used when responding. The outcomes of interest concerned the effect of target position and its interaction with any of the other variables. The main effect of target position slightly exceeded the threshold for significance, $F(1, 45) = 3.80$, $MSE = .02$, $p = .058$, $\eta_p^2 = .08$, and neither two-way interaction was significant (*F*s < .25, *p*s > .650). However, there was a significant three-way interaction, $F(3, 135) = 9.92$, $MSE = .02$, $p < .001$, $\eta_p^2 = .18$.

To explore this interaction, separate 2 (phase: 1 vs. 2) x 2 (position of target during phase 1 and 2: same vs. different) repeated measures ANOVAs were conducted for each trial type. On exemplar old trials, phase was the only significant effect, $F(1, 45) = 21.95$, $MSE = .02$, $p < .001$, $\eta_p^2 = .33$. For the remaining trial types, the effect of phase was significant as well as the interaction[3]. Position alone was not significant for any trial type (all *F*s < 1.8, all *p*s > .190).

The three individual interactions were therefore explored with Holm-Šidàk corrected paired-samples *t*-tests. On exemplar switch trials, phase 1 accuracy was higher for same ($M = .81$) than different ($M = .74$) trials, $t(45) = 2.14$, $p = .038$, $d = .33$, perhaps denoting some stimulus-specific differences that made the different trial set harder to remember during the initial recognition. However, during phase 2 performance was better on different ($M = .73$) than same ($M = .67$) trials, $t(45) = -2.33$, $p = .049$, $d = -.33$. State switch trials showed the reverse effect during phase 2 recognition, $t(45) = 2.63$, $p = .024$, $d = .38$, with higher accuracy on same ($M = .86$) than different ($M = .79$) trials. No difference was found during phase 1, $t(45) = -.40$, $p = .693$, $d = -.05$. The interaction on state old trials was primarily driven by clear time-dependent forgetting on different trials, with accuracy declining from 78.24% during the initial recognition test to 63.5% on the delayed recognition test, $t(45) = 6.16$, $p < .001$, $d = .86$. Comparison of phase 1 and 2 performance on same trials did not yield a significant result, $t(45) = .86$, $p = .394$, $d = .17$.

Discussion

Experiment 3 found both time-dependent forgetting (as observed in the previous experiments) and an interaction between phase and condition, first seen in Experiment 1. Overall, retrieval practice boosted phase 2 performance in comparison to the baseline condition, matching Experiment 1. The retrieval practice group also outperformed the restudy condition, showing that any effects of retrieval practice were not exclusively due to mere exposure. Retrieval practice was particularly beneficial for state switch trials, where accuracy during phase 2 was substantially higher than it had been during phase 1, despite the seven-day interval (a similar though less pronounced effect was observed in Experiment 2). Additionally, state switch performance following retrieval practice was notably higher than both restudy and baseline conditions. Yet on exemplar trials using an old foil, there was clear time-dependent forgetting in the retrieval practice condition (replicating Experiment 2 but not Experiment 1), and no reliable difference between baseline and retrieval practice groups.

Retrieval practice was beneficial on state old trials, where the retrieval practice group performed better than both restudy and baseline groups, despite evidence of time-dependent forgetting. Such an effect partially supported Experiment 1, where there was temporal forgetting for state information following retrieval practice, yet in Experiment 1 there was also no significant difference in phase 2 performance for retrieval practice and baseline conditions. In the exemplar switch condition, retrieval practice accuracy exceeded baseline, but not restudy, conditions. Indeed, restudy seemed to confer some advantages, but primarily for exemplar stimuli (during phase 2, the restudy group performed better than baseline for state switch and exemplar old trials).

In summary, the precise effect of retrieval practice depended on the specific information being tested, as well as the foil type. There was also evidence that the spatial location of the target and foil affected the retrieval practice condition – if the target and foil changed location over the interval, performance improved for exemplar switch trials, in comparison to when the target and foil remained in the same location. In contrast, on state switch and state old trials, it was preferable for the target and foil to remain in the same location. A full discussion of these data in relation to the wider study context is outlined below.

GENERAL DISCUSSION

Time-based forgetting is a well-established phenomenon and an important limitation in human memory. Its effects were clearly shown in the present study, with recognition performance declining over one week by approximately 11% in Experiment 1 and 10% in Experiment 3. Yet the primary aim of the present study was to determine whether the loss of detailed VLTM over time could be slowed or eliminated through retrieval practice. Overall, the data supported this idea – in Experiment 2, where retrieval practice was always given, average time-dependent forgetting was just 2.4%. In Experiments 1 and 3, the retrieval practice group performed better than the baseline condition after one week, and this effect was not simply due to mere exposure during initial testing. A restudy condition was added to Experiment 3 that allowed participants a second encoding opportunity for phase 2 stimuli, without retrieval practice. Whilst repeated encoding did boost recognition accuracy at phase 2, in comparison to the baseline group, performance in the restudy condition was still worse than the retrieval practice group, as shown in the analysis of the phase x condition interaction.

These findings are consistent with a testing effect (e.g. Karpicke, 2012; Roediger & Butler, 2011) and other studies showing that detailed memories can persist under certain circumstances (e.g. Sekeres et al., 2016). However, examination of higher-order interactions revealed that retrieval practice did not offer universal advantages and instead was highly dependent upon the trial type and foil type. Furthermore, different results emerged in Experiments 1 and 3, with retrieval practice helping exemplar but not state memory in the former study, and vice versa in the latter study (for old foil types). Closer examination of the method and results provides some insight into this discrepant result. In Experiment 1, state information was forgotten over time, yet there was still a trend towards better performance in the retrieval practice than baseline condition. Experiment 2 replicated time-dependent forgetting for state stimuli when both old and new foils were presented during phase 2, and the same effect emerged for state old stimuli in Experiment 3. In this final experiment, however, there was a reliable difference between conditions, with the retrieval practice group outperforming both the baseline and study conditions after one week. Retrieval practice may preserve state information to some degree, and this is supported by the exemplar switch condition. Here, a phase 1 exemplar trial was converted to a state trial during phase 2, and accuracy modestly *increased* over the delay in Experiment 2. There was evidence of time-dependent forgetting for exemplar switch stimuli in Experiment 3, but the retrieval practice condition outperformed the baseline group during phase 2.

Retrieval of an exemplar memory may preserve information about that object’s state, but mere exposure to a specific exemplar at phase 1 also seemed to have some limited benefit for remembering the object’s state at phase 2. In the restudy condition, accuracy was similar to the retrieval practice condition and marginally better than the baseline group on exemplar switch trials. The overall trends suggest that state detail is lost over one week, but this forgetting may be slowed through retrieval practice. Repeated exposure to an exemplar in a specific state has a less pronounced benefit, but might modestly alleviate information loss.

The way in which memory for exemplar information responded to retrieval practice differed across the experiments. Specifically, retrieval practice proved to be beneficial in Experiment 1 but not Experiment 3 (indeed, in Experiment 3 there was no difference between retrieval practice and baseline conditions on exemplar old stimuli during phase 2, which was the only case where retrieval practice did not significantly improve accuracy in relation to baseline). In the first experiment, the effect of retrieval practice may have been exaggerated by the type of foil used, as the exemplar target and foil often differed on state information too, providing more information through which the items could be distinguished during recognition. Experiment 1 also kept the target and foil items in the same position in both recognition tests, whereas Experiment 3 varied object placement. In Experiment 2, which varied object placement too, there was significant forgetting on exemplar trials that used the same foil during both phases, and this was more pronounced on trials where the target and foil had changed location. Experiment 3 did not replicate this spatial location effect on exemplar old trials, but the absence of target-foil placement does not necessarily mean it had no impact – for instance, participants may have realised that the position of the target can change and is therefore an unreliable memory cue. Additionally, the wider analysis of target position in Experiment 3 showed that it affected recognition on all remaining trial types.

The combined evidence highlights a loss of exemplar memory when a more careful design arrangement was used, yet in other cases exemplar information could be preserved over time. In Experiment 2, there was no loss of accuracy over the delay when a new foil was presented on exemplar trials at phase 2, or when a phase 1 state trial switched to an exemplar trial at phase 2. Experiment 3 replicated the latter effect, with performance on state switch trials jumping from 72% at phase 1 to 83% at phase 2 in the retrieval practice condition (the only case of facilitation in that experiment). This suggests that retrieval practice can preserve exemplar information over time, but the effect may be foil dependent. On old exemplar trials, participants may have experienced difficulty distinguishing the target from foil in the second test, as robust memories for both objects may have been formed during phase 1 recognition (though this problem might be avoidable when the recognition decision can be based on other information, such as target location or differences in state). Yet where interference from old foils can be avoided (such as on exemplar new and state switch trials), accuracy persists at a high level. This interpretation is supported by performance in the restudy condition of Experiment 3. In this condition, participants saw the target twice during the picture presentation, but they were only exposed to the specific target/foil pairing during the phase 2 test. On "exemplar old" trials, then, the restudy group avoided exposure to the foil during phase 1, and accuracy exceeded both the retrieval practice and baseline conditions, plus time-dependent forgetting was absent. A similar effect was shown on state switch trials, where the restudy condition performed better than the baseline condition during phase 2, when exemplar memory was tested.

Existing theories of the testing effect may prove useful in understanding the present results. Some accounts propose that a retrieval attempt effectively elaborates the representation, for example by adding mediating information (e.g. Carpenter, 2009) or extra semantic features (e.g. Verkoeijen, Bouwmeester, & Camp, 2012). Other explanations suggest that retrieval practice creates additional retrieval routes through which the memory can be accessed in the future (Rawson, Vaughn, & Carpenter, 2015). Both of these ideas can explain the overall benefit of retrieval practice, yet it is unclear why the precise effect of retrieval was dependent upon the trial and foil type. In addition, these theories struggle to account for cases where retrieval practice did not improve performance over the restudy group.

Another theory, known as the bifurcation model (Kornell, Bjork, & Garcia, 2011; see also Rowland, 2014), does allow a second encoding opportunity to strengthen a memory. This would explain the heightened accuracy in the restudy compared to baseline conditions, although there is an expectation that retrieval practice will be most beneficial (at least if the memory is correctly retrieved). Consequently, some of the present study's specific findings are difficult to reconcile with this theory. For example, restudy boosted recognition accuracy at phase 2 for exemplar information, but had less success in preserving state information. It should be stressed that most of the existing accounts of the testing effect are intended to understand the benefits of retrieval practice when recalling verbal stimuli. When recognition memory is involved for visual stimuli, especially with a 2AFC task, the foils used may strongly determine the nature of any retrieval practice effect, as well as the type of information being remembered (state or exemplar).

The present findings also pose problems for some explanations of time-dependent forgetting. Classic forgetting theories based on decay and interference processes would expect accuracy to decline over one week, but some previously retrieved or re-encoded memories seemed resistant to decay and interference. Indeed, in some rare cases phase 2 performance exceeded that at phase 1 – such an effect is striking, given that the interval lasted seven days. Models based on context-shifts may be better suited to explaining the retrieval practice effect, as theories incorporating a role for contextual factors typically expect the reinstatement of contextual information to improve retrieval (Unsworth et al., 2012). If the target and foil pairing is thought of as a form of context, the retrieval practice condition have that context reinstated during phase 2, which should boost recognition accuracy. But as with the theories of the testing effect, this interpretation struggles to explain situations in which retrieval practice was not beneficial.

Other theories of forgetting have relevance to the present results. For example, fuzzy-trace theory states that retrieval of verbatim traces allows vivid remembering, and it anticipates repetition to be beneficial (Brainerd & Reyna, 2002). The initial retrieval attempt may have preserved the verbatim trace used to support VLTM, and the repetition of stimuli provided by restudy conferred a more minor benefit. Similarly, the trace transformation model allows both the detailed and schematised representations of an event to exist simultaneously (Winocur & Moscovitch, 2011), and the retrieval environment influences the type of memory that is involved (Winocur et al., 2010). As such, it may be that retrieving an object at an early stage allows the detailed representation to persist over a one-week delay. As already noted, however, the precise effect of retrieval practice and restudy was dependent upon the trial and foil type, meaning that the more specific effects reported here remain difficult to explain.

The other major finding emerging from the present study concerned the rate at which information was lost from memory, with the results indicating more rapid forgetting for state-level than exemplar-level information. A larger loss of accuracy for state than exemplar memory was found in the retrieval practice condition of Experiment 1, and an effect of time occurred for state but not exemplar trials in Experiment 2. In Experiment 3, there was a greater overall decline in accuracy on state old than exemplar old trials, and on exemplar switch than state switch trials. Both effects were suggestive of increased difficulty in retaining information about an object's state over time. These findings are compatible with the view that specific properties of a memory may be lost independently and in different ways (e.g. Brady et al., 2013; Sekeres et al., 2016), rather than being forgotten in an all-or-none manner. However, Andermane and Bowers (2015) reported similar forgetting rates for state and exemplar memory. This could have been influenced by methodological differences, as Andermane and Bowers (2015) tested independent groups of participants at each phase and presented many more stimuli and recognition trials. Even so, differential loss of visual information has been reported before, with Brady et al. (2013) recording independent forgetting of state and exemplar memory for the same object over 3 days, and faster forgetting of memory for colour than state information. It should be noted that not all conditions conformed to the trend of faster forgetting of state information – in Experiment 1, there was more rapid forgetting of exemplar memory in the baseline condition, although this effect was not replicated in Experiments 2 and 3, and may have been produced by picture-specific effects. Indeed, at an item-level, some stimuli seemed much more difficult to retain over time than others.

Given these results, theories of VLTM may benefit from distinguishing between memory for state and exemplar information. Whilst exemplar memory is incorporated into formal theories such as the exemplar-based random walk model (Nosofsky & Palmeri, 1997), there do not appear to be similar theoretical frameworks for the retention of state information and it remains poorly understood. Subsequent studies would also benefit from employing a standardised set of stimuli, which was a particular limitation of Experiment 2 (specifically, the three recognition trial types were not always remembered equivalently in phase 1, although the primary focus in that experiment was to examine changes to each trial type over one week). Whilst the present study always ensured that the target and foil differed in either state or exemplar detail, the relationship between the target and foil may be important. Specifically, the difference between the target and foil was not standardised, and hence it varied from trial to trial. Participants may have been able to make use of this information, which may sometimes lead to an advantage. Indeed, both Andermane and Bowers (2015) and Cunningham, Yassa, and Egeth (2015) reported that recognition accuracy is much worse if a Yes-No or Old-New task is used, rather than the standard 2AFC test, and there is scope to further manipulate the way in which recognition memory is tested in future studies.

CONCLUSION

Time-based forgetting is a known limitation with human memory and many theories expect representations – especially those containing precise details – to be lost over a delay. Contrary to this view, the present study showed that an initial retrieval attempt preserved some finely resolved representations over seven days, even when the retrieval attempt was itself followed by a long delay. However, this effect was dependent on both the information being tested and the foil type, meaning that the complex set of results emerging in these three experiments is difficult to fully reconcile with existing theories of retrieval practice and time-dependent forgetting.

Word count: 11,870

REFERENCES

Andermane, N., & Bowers, J. S. (2015). Detailed and gist-like visual memories are forgotten at a similar rate over the course of a week. *Psychonomic Bulletin & Review, 22*, 1358-1363. 10.3758/s13423-014-0736-9

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *PNAS, 105*, 14325-14329. doi:10.1073/pnas.0803390105

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2013). Real-world objects are not represented as bound units: Independent forgetting of different object details from visual memory. *Journal of Experimental Psychology: General, 14*, 791-808. doi:10.1037/a0029649

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*, 164-169. doi:10.1111/1467-8721.00192

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS ONE, 5*(5), e10773. doi:10.1371/journal.pone.0010773

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1563-1569. doi:10.1037/a0017021

Cunningham, C. A., Yassa, M. A., & Egeth, H. E. (2015). Massive memory revisited: Limitations on storage capacity for object details in visual long-term memory. *Learning & Memory, 22*, 563-566. doi: 10.1101/lm.039404.115

Ebbinghaus, H. (1885). *Uber das gedächtnis*. Leipzig: Dunker & Humbolt.

Guerin, S. A., Robbins, C. A., Gilmore, A. W., & Schachter, D. L. (2012). Retrieval failure contributes to gist-based false recognition. *Journal of Memory and Language, 66*, 68-78. doi:10.1016/j.jml.2011.07.002

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences, 17*, 111-120. doi:10.1016/j.tics.2013.01.001.

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 396-411. doi:10.1037/0278-7393.31.3.396

Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principles. *Canadian Journal of Experimental Psychology, 63,* 124-138. doi:10.1037/a0014164.

Kang, S. H. K. (2010). Enhancing visuospatial learning: The benefit of retrieval practice. *Memory & Cognition, 38*, 1009-1017. doi:10.3758/MC.38.8.1009

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*, 157-163. doi:10.1177/0963721412443552

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968. doi:10.1126/science.1152408

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General, 139*, 558-578. doi:10.1037/a0019165

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language, 65*, 85-97. doi:10.1016/j.jml.2011.04.002

McKeown, D., Holt, J., Delvenne, J-F., Smith, A., & Griffiths, B. (2014). Active versus passive maintenance of visual nonverbal memory. *Psychonomic Bulletin & Review, 21*, 1041-1047. doi:10.3758/s13423-013-0574-1

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*, 266-300. doi:10.1037/0033-295X.104.2.266

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198. doi:10.1037/h0049234

Rawson, K. A., Vaughn, K. E., & Carpenter, S. K. (2015). Does the benefit of testing depend on lag, and if so, why? Evaluating the elaborative retrieval hypothesis. *Memory & Cognition, 43*, 619-633. doi:10.3758/s13421-014-0477-z.

Rickard, T. C., & Pan, S. C. (2017). A dual memory theory of the testing effect. *Psychonomic Bulletin & Review*. Advance online publication. doi:10.3758/s13423-017-1298-4.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15,* 20-27. doi:10.1016/j.tics.2010.09.003

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140,* 1432-1463. doi:10.1037/a0037559

Schachter, S. (1999). The seven sins of memory. *American Psychologist, 54*, 182-203. doi:10.1037//0003-066X.54.3.182

Schachter, S. (2001). *The seven sins of memory: How the mind forgets and remembers*. New York: Houghton Mifflin.

Sekeres, M. J., Bonasia, K., St-Laurent, M., Pishdadian, S., Winocur, G., Grady, C., & Moscovitch, M. (2016). Recovering and preventing loss of detailed memory: differential rates of forgetting for detail types in episodic memory. *Learning & Memory, 23*, 72-82. doi:10.1101/lm.039057.115

Squire, L. R. (2006). Lost forever or temporarily misplaced? The long debate about the nature of memory impairment. *Learning & Memory, 13*, 522-529. doi:10.1101/lm.310306

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterises flashbulb memories. *Psychological Science*, *14*, 455-461. doi:10.1111/1467-9280.02453

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2012). Dynamics of context-dependent recall: An examination of internal and external context change. *Journal of Memory and Language, 66*, 1-16. doi:10.1016/j.jml.2011.05.001

Verkoeijen, P. P. J. L., Bouwmeester, S., & Camp, G. (2012). A short-term testing effect in cross-language recognition. *Psychological Science, 23*, 567-571. doi:10.1177/0956797611435132

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society, 17*, 1-15. doi:10.1017/S1355617711000683

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal-neocortical interactions. *Neuropsychologia, 48*, 2339-2356. doi:10.1016/j.neuropsychologia.2010.04.016

Wixted, T. J. (2005). A theory about why we forget what we once knew. *Current Directions in Psychological Science, 14*, 6-9. doi:10.1111/j.0963-7214.2005.00324.x

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*, 409-415. doi:10.1111/j.1467-9280.1991.tb00175.x

FIGURE CAPTIONS

Figure 1. Example recognition trials from Brady et al. (2008). The pairings show target and foil items for exemplar trials (left column) and state trials (right column).

Figure 2. Mean proportion of correct responses for state (panel A) and exemplar (panel B) trials for immediate (phase 1) and delayed (phase 2) recognition. The dashed line shows the baseline condition and the solid line shows the retrieval practice condition. Error bars show 95% *CI*s calculated using Jarmasz and Hollands' (2009) equation for a mixed interaction.

Figure 3. Example recognition trials in Experiment 2. For phase 2 state and exemplar conditions, sometimes the foil from phase 1 was used again (old trials) and at other times a new foil was presented (new trials). On switch trials, the condition type changed over the delay (i.e. state-to-exemplar or exemplar-to-state). The target and foil swapped positions at phase 2 for half of the trials.

Figure 4. Mean proportion of correct responses for state (panel A) and exemplar (panel B) trials for immediate (phase 1) and delayed (phase 2) recognition. Each line shows one of the three recognition trial types (see legend for details). Error bars depict 95% *CI*s calculated using Jarmasz and Hollands' (2009) equation for a repeated measures interaction.

Figure 1.

Figure 2.

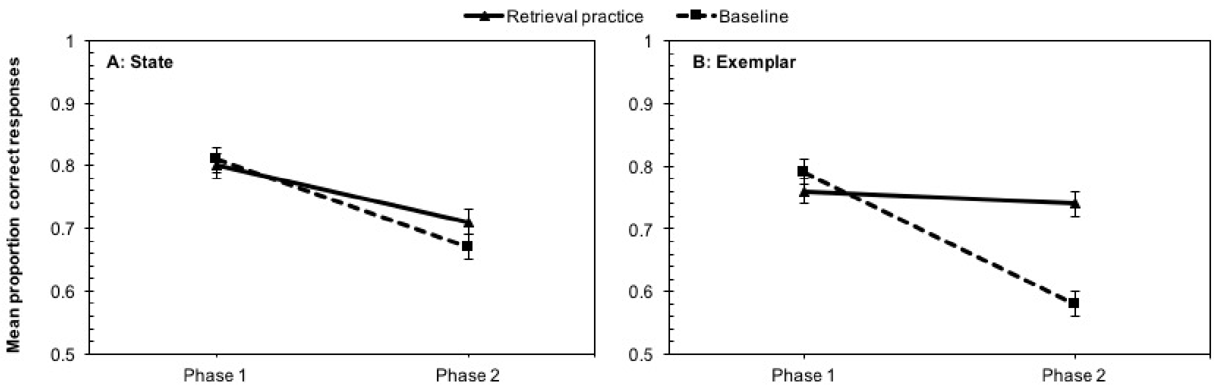

Retrieval practice
Baseline
A: State
B: Exemplar
Mean proportion correct responses
1
0.9
0.8
0.7
0.6
0.5
Phase 1
Phase 2

Figure 3.

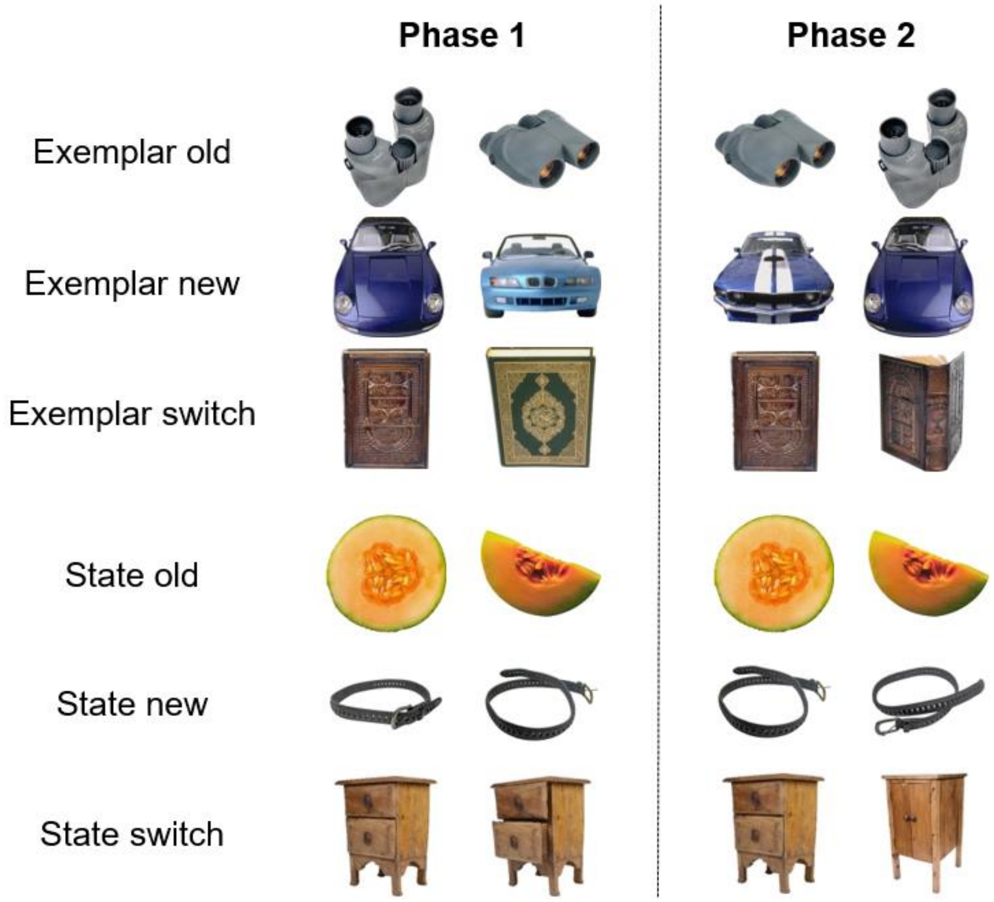
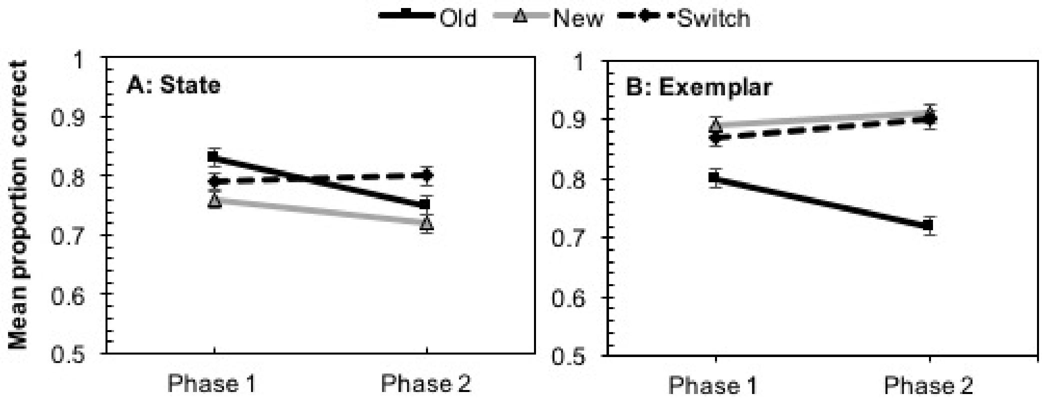

Phase 1
Phase 2
Exemplar old
Exemplar new
Exemplar switch
State old
State new
State switch

Figure 4.


Old
New
Switch
Mean proportion correct
A: State
B: Exemplar
1
0.9
0.8
0.7
0.6
0.5
Phase 1
Phase 2

TABLES

Table 1

Mean (and Standard Deviations) for Proportion of Correct Responses on Exemplar Old Trials According to Target Position During Phase 1 and 2

Target position in Phase 1 and 2	**Phase 1**	**Phase 2**
Same	0.84 (0.17)	0.78 (0.20)
Different	0.77 (0.18)	0.66 (0.18)

Table 2

Means (and Standard Deviations) for Proportion of Correct Responses in Experiment 3 According to Phase and Condition

Condition	Phase 1	Phase 2
Retrieval practice	0.75 (0.13)	0.72 (0.09)
Restudy	0.77 (0.12)	0.68 (0.09)
Baseline	0.80 (0.11)	0.62 (0.08)

Table 3

Time-Dependent Forgetting for each Trial Type and Condition in Experiment 3

Condition	Trial Type	Phase 1	Phase 2	Forgetting Rate	*t*	*p*	*d*
Retrieval practice	Exemplar old	*M* = 0.76, *SD* = 0.16	*M* = 0.68, *SD* = 0.15	0.08	3.47	.002	0.52
	Exemplar switch	*M* = 0.77, *SD* = 0.17	*M* = 0.70, *SD* = 0.12	0.07	2.98	.005	0.43
	State old	*M* = 0.76, *SD* = 0.14	*M* = 0.67, *SD* = 0.14	0.09	4.90	<.001	0.77
	State switch	*M* = 0.72, *SD* = 0.17	*M* = 0.83, *SD* = 0.10	+0.11	-5.57	<.001	-0.94
Restudy	Exemplar old	*M* = 0.73, *SD* = 0.15	*M* = 0.73, *SD* = 0.13	0	-.09	.932	0
	Exemplar switch	*M* = 0.83, *SD* = 0.15	*M* = 0.67, *SD* = 0.12	0.16	6.05	<.001	0.95
	State old	*M* = 0.77, *SD* = 0.12	*M* = 0.60, *SD* = 0.15	0.17	7.97	<.001	1.29
	State switch	*M* = 0.75, *SD* = 0.14	*M* = 0.71, *SD* = 0.14	0.04	2.37	.045	0.34
Baseline	Exemplar old	*M* = 0.76, *SD* = 0.17	*M* = 0.64, *SD* = 0.12	0.12	4.99	<.001	0.69
	Exemplar switch	*M* = 0.87, *SD* = 0.14	*M* = 0.62, *SD* = 0.12	0.15	11.93	<.001	1.59
	State old	*M* = 0.77, *SD* = 0.14	*M* = 0.57, *SD* = 0.13	0.20	9.05	<.001	1.16
	State switch	*M* = 0.79, *SD* = 0.12	*M* = 0.64, *SD* = 0.12	0.15	7.24	<.001	1.00

FOOTNOTES

---

[1] Participants were inadvertently shown one of the state objects twice during the picture presentation. Two groups were tested on this item during phase 1, and one of these groups was tested on this item again during phase 2. This image was replaced before the recognition test in the other two groups. Running the analysis without this trial did not affect the ANOVA output, nor did limiting the analysis to the two unaffected groups. The reported analysis therefore includes this object.

[2] Retrieval practice condition: Phase, $F(1, 45) = 6.30$, $MSE = .02$, $p = .016$, $\eta_p^2 = .12$; trial type, $F(3, 135) = 4.66$, $MSE = .01$, $p = .004$, $\eta_p^2 = .09$, phase x trial type, $F(3, 135) = 22.91$, $MSE = .01$, $p < .001$, $\eta_p^2 = .34$. Restudy condition: Phase, $F(1, 45) = 43.71$, $MSE = .02$, $p < .001$, $\eta_p^2 = .50$; trial type, $F(3, 135) = 5.59$, $MSE = .01$, $p = .001$, $\eta_p^2 = .12$, phase x trial type, $F(3, 135) =16.42$, $MSE = .01$, $p < .001$, $\eta_p^2 = .28$. Baseline condition: Phase, $F(1, 54) = 161.39$, $MSE = .02$, $p < .001$, $\eta_p^2 = .75$; trial type, $F(3, 135) = 10.10$, $MSE = .01$, $p < .001$, $\eta_p^2 = .16$, phase x trial type, $F(3, 135) = 8.16$, $MSE = .01$, $p < .001$, $\eta_p^2 = .13$.

[3] Exemplar switch: Phase, $F(1, 45) = 8.89$, $MSE = .03$, $p = .005$, $\eta_p^2= .17$; phase x position of target during phase 1 and 2, $F(1, 45) = 11.25$, $MSE = .02$, $p = .002$, $\eta_p^2= .20$. State old: Phase, $F(1, 45) = 24.64$, $MSE = .01$, $p < .001$, $\eta_p^2= .35$; phase x position of target during phase 1 and 2, $F(1, 45) = 11.08$, $MSE = .02$, $p = .002$, $\eta_p^2= .20$. State switch: Phase, $F(1, 45) = 30.93$, $MSE = .02$, $p < .001$, $\eta_p^2= .41$; phase x position of target during phase 1 and 2, $F(1, 45) = 4.41$, $MSE = .02$, $p = .041$, $\eta_p^2= .09$.