BRIEF REPORT

Pretesting boosts recognition, but not cued recall, of targets from unrelated word pairs

Tina Seabrooke [1] · Chris J. Mitchell [2] · Andy J. Wills [2] · Timothy J. Hollins [2]

Abstract

Attempting to retrieve the answer to a question on an initial test can improve memory for that answer on a subsequent test, relative to an equivalent study period. Such retrieval attempts can be beneficial even when they are unsuccessful, although this benefit is usually only seen with related word pairs. Three experiments examined the effects of pretesting for both related (e.g., *pond-frog*) and unrelated (e.g., *pillow-leaf*) word pairs on cued recall and target recognition. Pretesting improved subsequent cued recall performance for related but not for unrelated word pairs, relative to simply studying the word pairs. Tests of target recognition, by contrast, revealed benefits of pretesting for memory of targets from both related and unrelated word pairs. These data challenge popular theories that suggest that the pretesting effect depends on partial activation of the target during the pretesting phase.

Keywords Memory · Learning · Testing · Pretesting

Introduction

Tests have been hailed in recent years as effective and efficient studying tools (e.g., McDaniel, Anderson, Derbish, & Morrisette, 2007). It is well established that retrieving information from memory on an initial test can improve memory on later tests (Roediger & Karpicke, 2006). An important question, however, is *when* tests should be introduced. Traditional learning theorists argued that testing prematurely would be counterproductive, because failed tests could create confusion (Skinner, 1958; Terrace, 1963). Others strongly disagree, however, arguing that waiting until errors can be safely avoided wastes valuable study time (e.g., Kornell & Vaughn, 2016; Metcalfe, 2017). These researchers cite carefully controlled experiments that suggest that, just like successful tests, failed tests can be beneficial.

Kornell, Hays, and Bjork (2009) developed a pretesting procedure to examine the effects of failed tests on memory. Their participants first studied a list of weakly associated cue-target word pairs (e.g., *pond-frog*), by either studying each pair for the trial duration (Read condition), or by guessing the target for each cue before it was revealed (Generate condition). Since the word pairs were only weakly related, the participants' guesses on Generate trials were usually wrong. Nevertheless, the participants recalled more Generate targets than Read targets in a subsequent cued-recall test, and this pattern remained even when only trials involving incorrect guesses were analysed. Kornell et al. (2009) therefore showed that failed tests were constructive – an effect known as the *pretesting effect*.

Several theories have been proposed to explain the pretesting effect. Search set theory, for example, suggests that the process of generating guesses at encoding covertly activates a semantic network of related concepts, including the correct target. This prior activation is then suggested to improve the encoding of that target when it is subsequently revealed (Grimaldi & Karpicke, 2012; Hays, Kornell, & Bjork, 2013; Kornell et al., 2009; Zawadzka & Hanczakowski, 2018). Thus, search set theory emphasises the importance of the target already being partially activated when it is revealed. A corollary of this claim is that the locus of the effect is on the target itself, rather than the cue-target association.

Support for search set theory comes from the finding that pretesting improves subsequent cued recall for related (e.g., *tide-beach*) but not unrelated (e.g., *pillow-leaf*) word pairs (Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012;

✉ Tina Seabrooke
   t.seabrooke@soton.ac.uk

[1] Department of Psychology, University of Southampton, Southampton, UK

[2] Faculty of Health and Human Sciences, Psychology, University of Plymouth, Plymouth, UK

Knight, Ball, Brewer, DeWitt, & Marsh, 2012). Search set theory predicts this result because unrelated targets should not be activated in the search set for a cue, and so should not benefit from prior activation. More recently, Zawadzka and Hanczakowski (2018) used homograph cues that had two possible targets (e.g., *arms-hug* and *arms-nuclear*), although participants only ever saw one target. Pretesting was only beneficial when the participants' guesses were related to the target (e.g., guess *shoulder* for the pair *arms-hug*). This finding suggests that a pre-existing semantic relationship is not sufficient; the participant must also anticipate the correct relationship when guessing. Search set theory also predicts this result, because it is only when the correct relationship is assumed that the correct target might be activated.

Despite the findings above, support for search set theory is not universal. In particular, several recent experiments found that incorrectly guessing the definitions of novel English words or foreign vocabulary improved subsequent recognition of those definitions, relative to just studying the definitions (Potts, Davies, & Shanks, 2019; Potts & Shanks, 2014; Seabrooke, Hollins, Kent, Wills, & Mitchell, 2019a). The cues were unfamiliar words in these experiments, and so the targets were very unlikely to be part of any search set.

In summary, several studies on the one hand have supported search set theory by showing beneficial effects of guessing for subsequent *cued recall* of targets that were related to the cue, but not targets that were unrelated to the cue. In contrast, a handful of more recent studies have shown benefits of guessing for unrelated materials (using novel cues that have no associates), when tested using *target recognition* tests. Thus, there is an apparent conflict between the findings for recall and recognition, but this comparison is confounded with the materials used. Here we address this confound in three experiments.

The present experiments tested search set theory by examining the effects of pretesting on both target recognition and cued recall, using the same materials. In these experiments, the cues were all familiar words (e.g., *pond*) that would likely have many pre-existing semantic associates. The cues could be either related or unrelated to the targets. During the initial encoding phase, participants either studied the intact word pairs for the full trial duration (Read condition), or were presented with the cue and had to guess the target before the correct target was revealed (Generate condition). Memory for the targets from each pair was then assessed via a cued recall (Experiment 1) or target recognition (Experiment 2) test. In Experiment 3, both tests were administered in a single experiment. Search set theory predicts that the Generate condition should only improve memory for related word pairs, because it is only in the related condition that the target should form part of the search set for the cue. Crucially, this pattern should be seen in both recognition and recall. If the effects described above using novel cues (unfamiliar or foreign

words) are mediated by a similar mechanism to those seen with familiar cues, however, then we might expect guessing attempts to improve recognition of targets from both related and unrelated word pairs.

Experiment 1

Method

Participants Thirty participants (23 females; age: 18–36 years, $M = 22.37$, $SEM = 0.77$) were recruited from the University of Plymouth. This sample size provides 85% power to detect a medium-sized interaction, based on our prediction that the Generate > Read effect would be larger for related pairs than unrelated pairs. The University of Plymouth Psychology Ethics Committee approved all reported experiments.

Apparatus and materials The experiment was programmed in E-Prime 2.0 and was presented on a 22-in. computer monitor. Thirty-two related and 32 unrelated word pairs were selected from Nelson, McEvoy, and Schreiber's (1998) norms. The related pairs had forward associative strengths between 0.050 and 0.054, and the unrelated word pairs had no pre-existing associations. Allocation of word pairs to conditions was randomised for each participant.

Procedure The experiment consisted of encoding, distractor and test phases. At encoding, related (e.g., *bowl-plate*) and unrelated (e.g., *band-rash*) word pairs were presented. Figure 1a depicts the trial structure (for all experiments). On
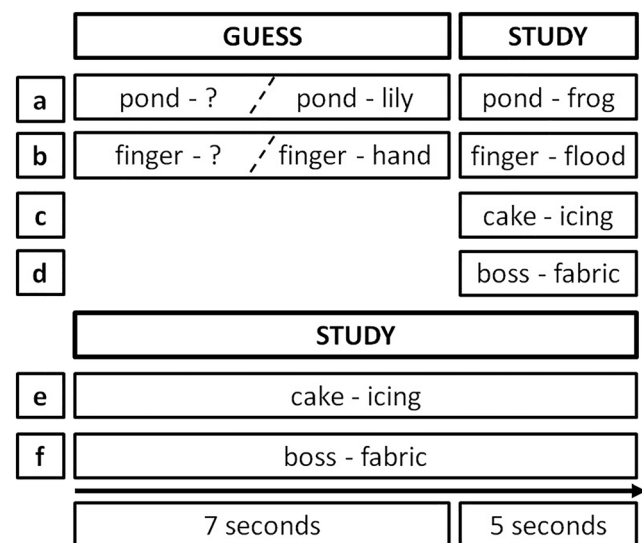

Fig. 1 Trial layout for Experiments 1, 2 and 3, for related and unrelated word pairs. Rows A and B depict the structure of Generate trials in all experiments. Rows C and D show the structure of Read trials in Experiments 1 and 2, and rows E and F show the structure of Read trials in Experiment 3

Read trials, participants studied the complete word pair for 5 s. On Generate trials, the cue (e.g., *bowl*) was first presented alone, and the participants had 7 s to guess the target (*plate*). Their guesses appeared on-screen as they typed, and they could use the backspace key to change their answer. After the 7 s had elapsed, the complete word pair was presented for 5 s. Four practice trials (two Read and two Generate, each with one related and one unrelated word pair) were administered first. The main encoding phase then consisted of 30 Read and 30 Generate trials, with 15 related and 15 unrelated word pairs presented in each. The trials were randomly intermixed and were separated by 500-ms intervals throughout all encoding and test phases in all reported experiments.

The distractor task between encoding and test lasted approximately 30 s, during which time participants evaluated the accuracy of simple mathematical statements (e.g., *(6 x 2) - 2 = 10*) by means of a button-press.

At test, the cues from each word pair were presented individually and the participants had to type in the corresponding target (or guess if they could not remember). Responses were not time-limited at test in any of the reported experiments. Four practice trials, using the cues from the practice encoding trials, were administered first, followed by the 60 main encoding phase cues.

Results and discussion

On average, the participants correctly guessed 5.78% (*SEM* = 0.94%) of targets from related word pairs at encoding. These word pairs were removed from further analysis. No targets from unrelated pairs were guessed.

Figure 2a shows the mean percentage of correctly recalled targets in the cued-recall test. An encoding condition (Generate, Read) × relatedness condition (related, unrelated) ANOVA revealed main effects of encoding condition, $F(1, 29) = 4.59$, $p = .04$, $\eta_g^2 = .02$, and relatedness condition, $F(1, 29) = 210.84$, $p < .001$, $\eta_g^2 = .54$, and a significant interaction, $F(1, 29) = 17.29$, $p < .001$, $\eta_g^2 = .05$. Relative to the Read condition, the Generate condition enhanced cued recall performance for related, $t(29) = 3.47$, $p = .002$, $d_z = 0.63$, $BF_{10} = 21.39$, but not unrelated, $t(29) = 1.02$, $p = .32$, $d_z = 0.19$, $BF_{10} = 0.31$, pairs (all Bayes Factors were calculated using Morey, Rouder, and Jamil's *BayesFactor* package [2015, version 0.9.12.4.2]). These results are consistent with previous demonstrations that pretesting only benefits cued recall for related word pairs (e.g., Grimaldi & Karpicke, 2012).

Experiment 2

Replicating prior research, Experiment 1 found that pretesting improved subsequent cued recall of targets from related, but not unrelated, word pairs. Experiment 2 sought to test search

set theory's prediction that the same pattern would be seen in any test of memory for the targets, using an old–new recognition test.

Method

Participants, apparatus and materials Thirty participants (25 females; age: 18–51 years, $M = 20.80$, $SEM = 1.07$) took part (with one participant replaced because of a computer failure). The sample size was based on a power analysis that used the same criteria as Experiment 1. Sixty-two related and 62 unrelated word pairs were selected as in Experiment 1. Thirty pairs from each relatedness condition were presented at encoding, with 15 each presented on Generate and Read trials. For each relatedness condition, two additional pairs were used for practice trials, and the targets from the remaining 30 pairs served as foils. Other aspects were as in Experiment 1.

Procedure The procedure was the same as in Experiment 1, except that the final test consisted of a target recognition test. Here, participants were asked to determine whether a target word was presented at encoding by clicking Yes/No buttons on the screen, using the mouse. Four practice trials were administered first (using targets from the practice encoding trials), followed by the 60 encoding phase targets and 60 novel foils.

Results and discussion

On average, the participants guessed 4.22% (*SEM* = 0.87%) of targets from related word pairs at encoding. These word pairs were removed from further analysis. No targets from the unrelated condition were guessed.

On average, 89.33% (*SEM* = 1.49%) of foils were correctly identified as novel in the target recognition test. Figure 2b shows the percentage of correct responses (hits) to the remaining (old) targets. There were significant main effects of encoding condition, $F(1, 29) = 81.15$, $p < .001$, $\eta_g^2 = .29$, and relatedness condition, $F(1, 29) = 11.80$, $p = .002$, $\eta_g^2 = .05$, but no significant interaction, $F(1, 29) = 1.28$, $p = .27$, $\eta_g^2 = .006$. The Generate condition improved subsequent recognition of both related, $t(29) = 7.30$, $p < .001$, $d_z = 1.33$, $BF_{10} = 295552.50$, and unrelated, $t(29) = 5.85$, $p < .001$, $d_z = 1.07$, $BF_{10} = 7869.44$, word pairs. This finding – that guessing improved target recognition memory for related and unrelated pairs – contrasts with Experiment 1, where guessing only improved cued recall for related word pairs.

Experiment 3

Experiments 1 and 2 suggest that pretesting improves memory for unrelated pairs in tests of target recognition, but not cued
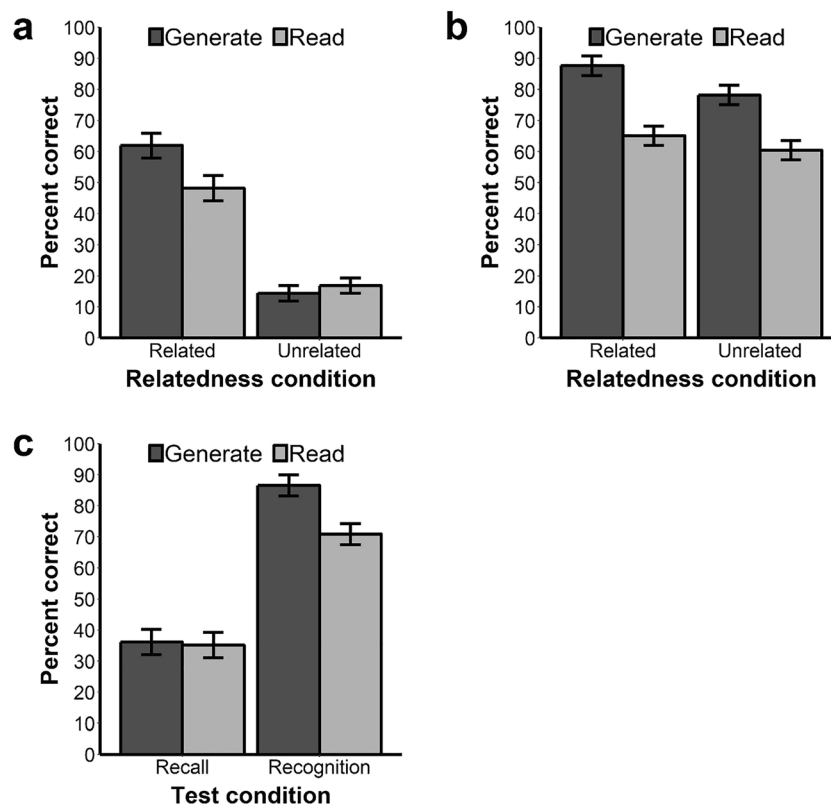
Fig. 2 Results from all experiments. (**a**) Experiment 1: Mean percentage of correctly recalled targets in the cued-recall test. (**b**) Experiment 2: Mean percentage of correct responses (hits) to old targets in the target recognition test. (**c**) Experiment 3: Mean percentage of correctly recalled targets in the cued-recall test, and mean percentage of hits to old targets in the target recognition test. Note that all tested word pairs were unrelated in Experiment 3. Error bars represent difference-adjusted within-subject 95% confidence intervals (Baguley, 2012)

recall. However, drawing strong conclusions across experiments is difficult, particularly given the low level of recall observed in Experiment 1. Experiment 3 therefore sought to directly compare recognition and recall of targets from unrelated pairs in a single study, and to improve the observed level of recall. Participants first studied related and unrelated word pairs at encoding. Memory for half of the unrelated pairs was then tested via a cued-recall test; the remaining unrelated pairs were tested via a target recognition test.

Two additional changes were made to the procedure in Experiment 3. The first was that we provided participants with the first letter of the targets during the cued-recall test (cf. Zawadzka & Hanczakowski, 2018), to improve overall recall performance. Additionally, in the previous experiments, the conditions were matched on target presentation time. The Generate trials were consequently much longer than the Read trials. In Experiment 3, we therefore matched the conditions on total trial time instead.

Method

Participants, apparatus and materials Thirty-six participants (19 males; age: 19–40 years, $M = 24.64$, $SEM = 0.95$) were recruited from the University of Plymouth or the University of Exeter for £4 each. The sample size provided 90% power to detect a medium-sized interaction effect.

The experiment was presented on a desktop or laptop PC (depending on experiment location). Thirty-eight unrelated and 26 related word pairs were selected for presentation, using the same criteria as the previous experiments. For each relatedness set, 24 pairs were presented at encoding, and a further two pairs on practice trials. The targets from the remaining 12 unrelated pairs served as foils in the target recognition test. Other aspects were as in Experiment 1.

Procedure The encoding and distractor phases were the same as the previous experiments, except that the Read trials at encoding lasted for 12 s (see Fig. 1), and the main encoding phase consisted of 48 trials (24 Generate and 24 Read). Related word pairs were presented on 12 trials in each encoding condition; unrelated pairs were presented on the remaining 12 trials. For each encoding condition (Generate/Read), six unrelated pairs were randomly allocated to the target recognition test; the remaining six unrelated pairs were allocated to the cued-recall test.

The test phase consisted of cued recall and target recognition trials, which followed the format of the test trials in Experiments 1 and 2, respectively, except we provided the

first letter of the target on cued-recall trials. Both tests assessed memory for just the unrelated word pairs. Two practice trials from each trial type were administered first. The main test consisted of 12 cued recall trials (six Generate and Read pairs each), intermixed with 24 target recognition trials (six Generate targets, six Read targets and 12 foils). Other aspects were as in the previous experiments.

Results and discussion

On average, the participants guessed 6.71% ($SEM = 1.14\%$) of targets from related word pairs at encoding. No targets from the unrelated condition were guessed. Since only unrelated pairs were presented at test, no pairs were removed from the test dataset.

On average, 87.04% ($SEM = 1.95\%$) of foils were correctly identified as novel in the target recognition test. Figure 2c shows the percentage of correct responses (hits) for remaining (old) targets in the recognition test, and correctly recalled targets in the cued-recall test. An encoding condition (Generate, Read) × test format (cued recall, target recognition) ANOVA revealed significant main effects of encoding condition, $F(1, 35) = 10.00$, $p = .003$, $\eta_g^2 = .03$, and test format, $F(1, 35) = 132.94$, $p < .001$, $\eta_g^2 = .43$, and a significant interaction, $F(1, 35) = 8.19$, $p = .007$, $\eta_g^2 = .02$. The Generate condition produced better subsequent target recognition than the Read condition, $t(35) = 4.75$, $p < .001$, $d_z = 0.79$, $BF_{10} = 658.85$. The encoding conditions did not differ in cued recall, $t(35) = 0.23$, $p = 0.82$, $d_z = 0.04$, $BF_{10} = 0.18$. The results therefore replicate the previous results; pretesting improved recognition but not cued recall of targets from unrelated word pairs.

General discussion

Three experiments revealed differential effects of pretesting on cued recall and target recognition. Relative to just studying word pairs, pretesting improved subsequent recall of targets from related, but not unrelated, word pairs. In tests of target recognition, by contrast, pretesting improved memory for targets from all pairs.

Although we successfully replicated the selective-recall pattern predicted by search set theory, the novel finding from these experiments is that pretesting improved target recognition for both related and unrelated pairs. This pattern is *inconsistent* with the prediction from search set theory, which predicts that guessing should not lead to greater activation of an unrelated target. Instead, this latter result accords with previous vocabulary learning experiments that observed benefits of generating errors in target recognition tests (Potts et al., 2019; Seabrooke et al., 2019a). Together, these results suggest that search set theory has limitations in explaining the pretesting effect.

Attentional accounts provide an alternative to search set theory (e.g., Potts & Shanks, 2014; Zawadzka & Hanczakowski, 2018). Such accounts suggest that pretesting increases attention to the subsequent feedback (the target), relative to study alone. Enhanced attention at encoding should produce a richer memory trace and should, therefore, improve subsequent recognition. There are several ways in which pretesting could boost attention to corrective feedback. One possibility is that incorrect guesses produce surprise when corrective feedback is provided, which then boosts attention to the target through an error-correction mechanism (although see Zawadzka & Hanczakowski, 2018). Pretesting may alternatively improve attention to feedback by increasing motivation. Recent work has shown that, relative to study, pretesting increases self-reported curiosity (Potts et al., 2019) and motivation to learn an answer (Seabrooke, Mitchell, Wills, Waters, & Hollins, 2019b). For the present purposes, we do not distinguish between the motivational and error-correction accounts, but instead refer to them jointly as "attentional" theories of pretesting effects.

Attentional accounts predict that pretesting should improve recognition of targets from both related and unrelated pairs, because incorrect guesses that are followed by corrective feedback should trigger an error-correction mechanism and/or enhance motivation in both cases. Thus, unlike search set theory, attentional accounts readily predict the observed recognition benefits of pretesting for targets from both related and unrelated pairs.

Attentional accounts do, however, need to explain why pretesting improves *cued recall* of related pairs only. One possibility is that there is a separate retrieval process that boosts cued recall for just related pairs. Consider the example where a participant incorrectly guesses the target for the cue *pond*, either when the actual target is related (*frog*) or unrelated (*spanner*). In a subsequent recognition test, we know that participants will be more likely to correctly recognise both *frog* and *spanner* than targets that they just studied. This does not mean, however, that both targets will be equally accessible from the cue *pond* (Bjork & Bjork, 1992). What is accessible in cued recall is driven by many factors, including the prior semantic associates to the cue. When a participant has no episodic recollection of the target for a given cue, they may generate semantic associates of that cue. When faced with the cue *pond*, for instance, they may generate the semantically related item *frog*. By contrast, they will be much less likely to generate the unrelated item *spanner*. Moreover, if *frog* is generated, we know it is likely be recognised and thus participants will be even more likely to choose that correct answer. The same argument applies to Zawadzka and Hanczakowski's (2018) result where, for example, *arms-hug* only benefitted from guessing if the participant made the correct interpretation of *arms* (e.g., if they guessed *legs* rather than *missile*). On test, participants would be searching for semantic associates of

arms, and they may once again interpret the word *arms* as a body part rather than a weapon (and hence only body parts would be generated as candidate targets at test).

To conclude, the present experiments observed a benefit of pretesting over studying for related word pairs in both cued recall and target recognition. For unrelated pairs, generating errors also improved target recognition but not cued recall. The results add to a growing literature suggesting that generating errors improves memory for targets from unrelated word pairs. These findings are inconsistent with the idea that the benefits of generating errors arise during encoding, through a process of partial activation of the target when guessing. Instead, the results suggest that pretesting enhances attention to all targets, regardless of whether the cue and the target have a pre-existing semantic relationship. Further, they suggest that the differential effect of pretesting that is seen for related and unrelated word pairs in cued recall is a retrieval-based mechanism, rather than one operating during the pretesting phase. Here we propose one potential mechanism for this differential effect, namely the use of pre-existing semantic associations to support retrieval of related, but not unrelated, targets, but we do not rule out the possibility that alternative retrieval-based mechanisms may be able to account for this pattern.

Open practices statement All data are publicly archived at https://osf.io/ksajx/ (Experiment 1), https://osf.io/5hvqg/ (Experiment 2) and https://osf.io/nesxh/ (Experiment 3). The experiments were not preregistered.

References

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, *44*, 158–175. https://doi.org/10.3758/s13428-011-0123-7

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, *40*, 505–513. https://doi.org/10.3758/s13421-011-0174-0

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 290–296. https://doi.org/10.1037/a0028468

Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, *40*, 514–527. https://doi.org/10.3758/s13421-011-0167-z

Knight, J. B., Ball, B. H., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing unsuccessfully: A specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language*, *66*, 731–746. https://doi.org/10.1016/j.jml.2011.12.008

Kornell, N., Hays, M., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 989–998. https://doi.org/10.1037/a0015729

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of Learning and Motivation*, *65*, 183–215. https://doi.org/10.1016/bs.plm.2016.03.003

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*, 494–513. https://doi.org/10.1080/09541440701326154

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*, 465–489. https://doi.org/10.1007/BF01457248

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). Package 'BayesFactor.' Retrieved from https://richarddmorey.github.io/BayesFactor/

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved from http://w3.usf.edu/FreeAssociation/

Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: What is the locus of the effect? *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*, 1023–1041. https://doi.org/10.1037/xlm0000637

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, *143*, 644–667. https://doi.org/10.1017/CBO9781107415324.004

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

Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019a). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language*, *104*, 70–82. https://doi.org/10.1016/j.jml.2018.10.001

Seabrooke, T., Mitchell, C. J., Wills, A. J., Waters, J. L., & Hollins, T. J. (2019b). Selective effects of errorful generation on recognition memory: The role of motivation and surprise. *Memory*, *27*, 1250–1262. https://doi.org/10.1080/09658211.2019.1647247

Skinner, B. F. (1958). Teaching machines. *Science*, *128*, 969–977. https://doi.org/10.1109/TE.1959.4322064

Terrace, H. S. (1963). Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior*, *6*, 1–27. https://doi.org/10.1901/jeab.1963.6-1

Zawadzka, K., & Hanczakowski, M. (2018). Two routes to memory benefits of guessing. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000676