

Journal of Experimental Psychology: Learning, Memory, and Cognition

Effects of Varied and Constant Environmental Contexts on Acquisition and Retention

Steven M. Smith and Justin D. Handy

Online First Publication, May 5, 2014. <http://dx.doi.org/10.1037/xlm0000019>

CITATION

Smith, S. M., & Handy, J. D. (2014, May 5). Effects of Varied and Constant Environmental Contexts on Acquisition and Retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <http://dx.doi.org/10.1037/xlm0000019>

Effects of Varied and Constant Environmental Contexts on Acquisition and Retention

Steven M. Smith and Justin D. Handy
Texas A&M University

Four experiments examined the decontextualization of memories, the stage of learning in which memories can be recalled in the absence of episodic memory cues. Face–name pairs were studied with video-recorded environmental contexts in the background, and after 5 practice trials, recall of names associated with faces was tested in the absence of the original video context cues. In Experiments 1, 2, and 3, five retrieval practice trials for a pair occurred either always with the original video context (constant context condition) or with a new context on each trial (varied context condition). Final recall was tested either on the same day (Experiments 3 and 4) or 2 days later (Experiments 1 and 2), and either the recall test context for each face was a new (never seen) video (Experiment 1) or there was no context shown at test (Experiments 2, 3, and 4). In the first 3 experiments, acquisition was better for the constant context condition, but on the final recall test, performance was better for pairs learned under varied context conditions. In Experiment 4, which used multiple study trials rather than multiple retrieval practices during acquisition, no differences were found between the constant and varied context conditions, either for acquisition or for final retention. The results show that acquisition trials given in varied contexts can result in decontextualized memories, but only when acquisition involves retrieval practice, rather than simple restudy trials. These results are consistent with the new theory of disuse, but not with the theory of encoding variability.

Keywords: context, decontextualization, encoding variability

Tulving's distinction between episodic and semantic memories is that episodic memories are autobiographically referenced and remembered as experiences that occurred at some particular time and place, whereas semantic memories are known facts and may not be rooted in specifically experienced contexts (Tulving, 1972, 1983). When new knowledge is first experienced, memory of that information, having been experienced only once, may be highly context dependent (e.g., Godden & Baddeley, 1975; S. M. Smith, 1979; S. M. Smith, Glenberg, & Bjork, 1978; S. M. Smith & Vela, 2001). An example is memory of someone's name; after first meeting a person, recalling his or her name may involve retrieval of a singular contextualized memory, whereas after numerous encounters with that person in many contexts, memory of his or her name may involve semantic memory, that is, a known fact. Craik (1979) proposed "a continuum of representation, running from highly context-specific episodes at one extreme to abstract generalized knowledge at the other" (p. 451). How does knowledge progress from its initial dependence on specific episodic contextual cues to the type of knowledge that is independent of episodic contexts? In the present study, we examined the effects of repeated practice on newly learned face–name pairs; does varying the contexts of retrieval practice, relative to holding the contexts

constant, make memory for face–name pairs more *decontextualized*?

Not all semantic knowledge is episodically decontextualized; that is, if one learned a fact or a person's name in a single episodic context, then the fact or name might remain contextualized. Mandler's (1980) butcher-on-the-bus example describes someone who momentarily fails to recognize that a man seen on a bus is someone well known to the rememberer, a butcher normally encountered at the supermarket. Mandler and others (e.g., Gruppuso, Lindsay, & Masson, 2007) have focused research on the distinction between recognizing an item and recognizing its context, which may involve searching memory for the right context for a familiar face. In the present investigation, however, we did not study recognition or recollection of contexts, but rather the way that context can cue newly learned content, specifically, a name (such as Norman) associated with a face. The primary focus of the present investigation was on face-cued recall of newly learned names as a function of incidental environmental context cues.

Studies of Environmental Context-Dependent Memory

Studies that have examined the effects of manipulations of the incidental environmental contexts of events on memory of those events date back at least to S. Smith and Guthrie (1921) and Pan (1926) for studies involving human participants. Through the end of the 20th century, most reported studies of environmental context-dependent memory would present a large set of materials (such as lists of unrelated words) in a particular environment, testing memory for those materials either in the original environment (the reinstated condition) or in another one (see S. M. Smith & Vela, 2001, for a review). Not all attempts to examine environ-

Steven M. Smith and Justin D. Handy, Department of Psychology, Texas A&M University.

Correspondence concerning this article should be addressed to Steven M. Smith, Department of Psychology, Texas A&M University, College Station, TX 77843. E-mail: stevesmith@tamu.edu

mental context-dependent memory have found significant effects (e.g., Farnsworth, 1934; Fernandez & Glenberg, 1985; Saufley, Otaka, & Bavaresco, 1985; S. M. Smith et al., 1978), and overall, the reported effects have been significant, but modest, with an average effect size of $d = 0.28$ (S. M. Smith & Vela, 2001).

More recently, highly effective methods have been developed and used for examining robust effects of environmental context manipulations on memory (e.g., Hayes, Nadel, & Ryan, 2007; S. M. Smith, Handy, Angello, & Manzano, 2014; S. M. Smith & Manzano, 2010). The key differences between the more recent studies and most of the older ones are (a) newer studies typically use digitally represented environments, such as photos (e.g., Gruppuso et al., 2007; Hayes et al., 2007; Hockley, 2008) or video recordings (Jonker, Seli, & MacLeod, 2013; Staudigl & Hanslmayr, 2013; S. M. Smith & Manzano, 2010; S. M. Smith et al., 2014) of physical environments, rather than the actual physical cues, and (b) whereas older studies tended to use many memory targets in a single context, newer studies typically use few (e.g., one) targets per context and many different contexts.¹ The use of few memory targets per context cue (i.e., a small contextual fan) enhances the likelihood that the cue will evoke a memory target (e.g., S. M. Smith & Manzano, 2010), and the use of many context cues decreases the participant's ability to mentally reinstate context cues not provided by the experimenter (see S. M. Smith, 1979). Thus, this method gives the experimenter more power over the use of context cues, and it helps ensure that provided context cues will evoke retrieval of memory targets. In the present study, a version of S. M. Smith and Manzano's (2010) video context method was used to test contextualized and decontextualized face-name pairs.

Decontextualization

An operational definition of a decontextualized memory is a memory that is not adversely affected when associated context cues are removed, that is, a memory that is accessible even without the use of contextual cues. One experimental approach to this issue has been to repeat study material, either in a single context or in varied contexts, testing the resultant memory in the absence of context cues. For example, Gartman and Johnson (1972), testing free recall of a list of words, repeated some homographs within the list (e.g., *foot*) either biased by the same verbal context both times (e.g., *inch* and *meter* preceding the first presentation of *foot*, *mile* and *yard* preceding the second) or with two encoding contexts (e.g., *inch* and *meter* preceding the first presentation of *foot*, *arm* and *hand* preceding the second). On the free recall test, in which no context cues were provided, better recall was found for the variable encoding condition than for the constant context condition. Similarly, Experiment 1 of S. M. Smith et al.'s (1978) investigation, which manipulated environmental contexts during study repetitions, rather than verbal contexts, found that repetition of material in varied contexts, rather than the same context twice, resulted in better recall of the repeated material. In S. M. Smith et al.'s study, participants were tested in a new, never-seen environmental context, rather than either of the study contexts. These two investigations suggest that if no appropriate cues are provided at test, memory for repeated material is better if the repetitions occur in varied contexts, as opposed to holding context constant across repetitions.

The results of these varied and constant context studies have been interpreted in terms of encoding variability, which states that encoding an item repeated with the same context results in a single memory trace, whereas repetitions in varied contexts results in multiple memory traces. Decontextualized memories, according to this view, are more easily remembered because multiple different cues could provide access to memories, whereas contextualized knowledge depends on one particular episodic context cue. The problem with the single memory trace is analogous to "putting all your eggs in one basket"; that is, losing the single cue to a memory trace results in forgetting, whereas losing one of several possible cues nonetheless allows for other cues that could successfully evoke the memory trace.

A different approach to understanding decontextualized memories uses Hintzman's (1986) MINERVA 2 model, which Hintzman used to explain "schema abstraction," a schema representing a decontextualized type of knowledge. In Hintzman's theory, every experience is encoded in its own memory trace, including each repetition of a particular item of information, such as a word or an object. The process of memory retrieval in MINERVA 2 is based on Semon's (1909/1923) notion of *ecphory*, a resonance-like process that returns memory content to consciousness in the form of an echo (see Schacter, Eich, & Tulving, 1978). Cued retrieval begins with a memory probe, whose content coincides with that of all memory traces that contain the cue. A parallel summation of all such traces serves as the contents of the echo. In the echo, the similarities among the memory traces are amplified, representing an abstracted schema of the cued memory. If the incidental contexts of the summed traces in the echo are the same, they will be represented in the echo along with schematically relevant information. If the contexts are varied, however, the echo will more clearly represent the relevant content of the remembered material, which will be far more amplified in the echo than the associated episodic contexts. In face-name paired associates, for example, which were studied and recalled in the present investigation, the echo for pairs studied with varied contexts should more clearly represent the correct response than echoes for pairs studied with constant contexts. Thus, varied encoding contexts result in a more decontextualized memory in which the correct response can be distinguished from incidental contextual features (see also S. M. Smith, 1994).

Bjork and Bjork's (1992) new theory of disuse explains memory as a joint function of two parameters: storage strength, an index of an item's long-term memory potential, and retrieval strength, the ease of access to an item in memory at a given moment or in a particular situation. Because of the cue-dependent nature of memory retrieval, a name learned in a particular context will have greater retrieval strength if that context is reinstated, rather than altered. One of the assumptions of the new theory of disuse states that the act of retrieval results in an increase in the storage strength of the retrieved item, and the more difficult or involved the act of retrieval is, the more storage strength is gained by the act. That

¹ A notable exception to the older methods was Experiment 2 of the first environmental context-dependent study of humans that was reported in a scientific journal (Pan, 1926), in which face-name pairs were affixed to picture postcards of places in Chicago and tested with either the same pictured place or a different place. Pan's (1929) method was remarkably similar to the method reported in the present experiments.

memory benefits from more difficult retrieval has been referred to as the desirable difficulties principle (e.g., Finley, Benjamin, Hays, Bjork, & Kornell, 2011). The context variability effects reported by S. M. Smith et al. (1978) and Gartman and Johnson (1972) can be explained as the result of retrieval practice difficulty (Bjork & Bjork, 1992). When a repetition occurs in a new context, relative to a repetition in the original encoding context, it is more difficult to retrieve, as evidenced by poorer recall for material tested in new contexts (e.g., S. M. Smith et al., 2014). A greater increment in encoding strength should result from the more difficult retrieval repeated in a new context, according to this explanation. Although the results of those context variability experiments do not serve as a critical test of these two explanations, that is, new theory of disuse versus encoding variability, the present experiments provide a test of the two theories.

The Present Experiments

The present experiments were motivated by a combination of theoretical and methodological questions, both of which have implications for our understanding of decontextualized memories. Experiments 1, 2, and 3 tested predictions derived from the new theory of disuse with respect to the effects of varied contexts and constant contexts during acquisition of a face–name pair, and Experiment 4 tested an encoding variability account of the first three experiments. All four experiments examined the effects of constant and varied contexts during acquisition trials on final recall for newly learned face–name pairs. Participants initially studied a set of 20 face–name pairs, each pair shown superimposed over a 5-s video context scene. They were told that they would be tested for their memory of the names associated with the faces. Participants were not instructed to attend to the video contexts shown in the background; they were told simply that the videos were being shown because the experimenters were interested in how they affected performance. Following initial study in Experiments 1–3, participants were given five blocks of retrieval practice for the face–name pairs. The new theory of disuse predicted that relative to constant contexts, varied contexts across retrieval practice trials would impede measures of acquisition, because recall suffers when tested in new contexts. According to this theory, varied contexts should benefit retention, as measured by a final recall test. Experiment 4 resembled the first three experiments, except that acquisition trials involved only re-presentation of pairs, rather than retrieval practice. Whereas the encoding variability theory predicted that retention would be better for the varied context condition in all four experiments, the new theory of disuse predicted a varied context advantage in retention only for Experiments 1–3, which involved retrieval practice, but not Experiment 4, which involved only restudy trials during acquisition.

Experiment 1

Experiment 1 used the acquisition procedures described above, which involved initial study of face–name pairs, and five retrieval practice trials, done either in constant contexts (i.e., the same original study context on each trial) or in varied contexts (i.e., a new context on each retrieval practice trial). To decrease retrieval strength at test as much as possible, we used a 2-day retention interval, and each face cue was displayed at test with a new video

context in the background, that is, a video never seen by the participant until the final test. It was predicted that retrieval practice would be easier if the original context was reinstated on each retrieval practice trial, relative to showing a new context on each retrieval practice, which would make retrieval practice more difficult. That is, we predicted better performance during acquisition, as measured by success on retrieval practice trials, for the constant context condition, relative to the varied context condition, because reliable cues for retrieval were provided on every trial in the constant context condition, but not in the varied context condition. It was also predicted that final retention, tested without the presence of context cues, would show a reversal of the pattern seen during acquisition, with better performance in the varied context condition relative to the constant condition.

Method

Participants. Forty-five Texas A&M University undergraduate students participated in Experiment 1 in return for partial course credit. Participation was voluntary, and other options were available to earn equal credit. Participants self-enrolled in the experiment, and group assignment was based on time of testing. The number of participants in each experimental session depended upon the random number of participants who opted to enroll for a session, with a minimum of two and a maximum of 15 participants enrolling per session. There were 22 participants in the constant context condition, whereas 23 participants were assigned to the varied context condition.

Design and materials. Background context during retrieval practice (constant or varied) was manipulated between subjects. The dependent variables were (a) the proportion of names correctly recalled in each retrieval practice block and (b) the proportion of names correctly recalled on the final cued recall test 2 days later.

Twenty Caucasian male faces (all presented in color) were sampled from the Psychological Image Collection at Stirling (<http://pics.stir.ac.uk>). Each face was paired with a common first name (e.g., Michael, Daniel, Tyler) with the provision that each name began with a different letter. The face–name pairs were randomly paired with background movie scenes during initial study, retrieval practice, and the final cued recall test. One hundred and forty unique 5-s video clips were used. These color movie scenes depicted common, everyday events (e.g., walking through a bookstore, driving down a freeway, walking through a kitchen) complete with audio. Camtasia Studio video production software was used for stimuli presentation.

Procedure. Participants were tested as a group in front of a large projection screen. During the initial study phase, participants were told that they would study a series of faces paired with first names for a later memory test. Furthermore, they were told that each unique face–name pair would be presented over a different movie scene and that the experimenters were interested in how these background scenes would affect their memory. There was no mention that the background scenes would appear on the final memory test. After the instructional period, participants viewed all 20 face–name pairs and their associated movie scenes in a random order. Each study item remained on the screen for 5 s with a 2-s intertrial interval.

Following the study phase, participants were given instructions for completing the first retrieval practice block. For each test item, participants viewed a previously studied face and wrote down the corresponding name, guessing if necessary. Critically, each face was superimposed over a background movie scene. Participants in the constant context condition viewed each face paired with the same 5-s scene encountered during the initial study phase. In contrast, participants in the varied context condition were shown faces paired with new 5-s scenes. After writing down their responses, participants were given feedback in the form of the intact face–name pair along with its associated background scene for 5 s. During this time, participants were instructed to circle the name they wrote on their response form if it was incorrect. This test–feedback procedure continued until all 20 test items were practiced. Following this test format, participants completed five retrieval practice blocks with all 20 test items appearing in a randomized order in each block. Participants were excused for the day following the fifth retrieval practice block.

Upon returning 2 days later, participants were administered a final cued recall test. For each test item, they were shown a face superimposed over a brand-new 5-s movie scene. Their task was simply to write down the name that corresponded to each of the faces. Participants had 5 s to respond to each test item followed by a 2-s delay before presentation of the next test item. Similar to the retrieval practice blocks, each test item appeared in a randomized order.

Results

Acquisition. Unless specified otherwise, all statistical analyses used an alpha level of .05 for determining statistical significance. A 2×5 mixed analysis of variance (ANOVA) was computed, with context (constant and varied) as a between-subjects variable and retrieval practice trial (RP1, RP2, RP3, RP4, and RP5) as a repeated measure. The proportion of names correctly recalled on each retrieval practice trial served as the dependent measure. There was a significant main effect of retrieval practice trial, $F(4, 172) = 203.99$, $MSE = 2.68$, $p < .0001$. There was also significant main effect of context, $F(1, 43) = 36.92$, $MSE = 4.16$, $p < .0001$; retrieval practice performance was greater in the constant context condition ($M = 0.64$, $SD = 0.14$) than the varied context condition ($M = 0.37$, $SD = 0.14$). Additionally, the interaction between retrieval practice trial and context condition was significant, $F(4, 172) = 4.18$, $MSE = 0.055$, $p < .003$. To examine the interaction, pairwise comparisons between retrieval practices in the constant and varied context conditions were analyzed with a stricter rejection criterion of $p < .01$ to control for family-wise error associated with multiple comparisons. The analysis showed significant differences between constant and varied context conditions for all retrieval practice trials ($ps < .0001$; see Figure 1).

Retention and forgetting. For the final cued recall test, performance was measured along two dimensions: (a) retention on the final test and (b) forgetting occurring between the final retrieval practice block and the final test. First, a comparison of retention rates between the constant ($M = 0.43$, $SD = 0.25$) and varied ($M = 0.61$, $SD = 0.26$) context conditions revealed a significant difference in the proportion of face–name pairs correctly recalled, $t(43) = -2.29$, $p = .03$, Cohen's $d = 0.71$.

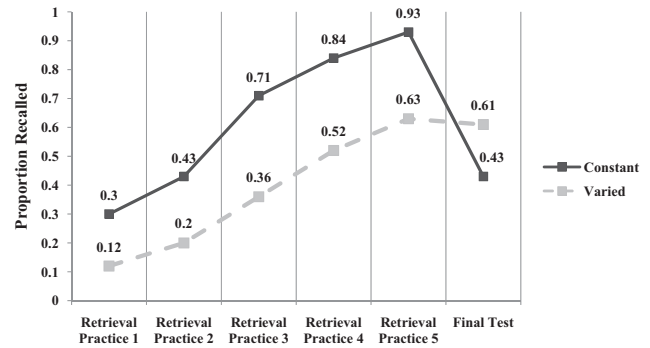


Figure 1. Proportion correctly recalled in Experiment 1 on each of five retrieval practice trials and on the final retention test as a function of context condition.

Forgetting was assessed by first considering the difference in the proportion of study items recalled in Retrieval Practice 5 versus the proportion of study items recalled in final cued recall test. Whereas forgetting rates were dramatic for participants in the constant context condition ($M = 0.50$, $SD = 0.24$), participants in the varied context condition ($M = 0.02$, $SD = 0.17$) displayed minimal forgetting between the two time intervals, $t(43) = 7.73$, $p < .001$, Cohen's $d = 2.34$. A complementary analysis evaluated the percentage of study items from Retrieval Practice 5 forgotten on the final test. Consistent with the previous analysis, the percentage of study items lost between the two time intervals was greater for the constant context condition ($M = 0.54$, $SD = 0.26$), with no losses in the varied context condition ($M = -0.05$, $SD = 0.49$), $t = 5.02$, $p < .0001$, Cohen's $d = 1.57$.

Discussion

As predicted by the new theory of disuse, we found that varying contexts during retrieval practice, relative to holding contexts constant across practice trials, impeded acquisition of face–name pairs but benefited retention, as measured by a final recall test. Recall on retrieval practice trials was better in the constant context condition than the varied context condition, showing significant differences between constant and varied context conditions for all retrieval practice trials. On the final retrieval practice trial, performance in the constant context condition was 25% greater than performance in the varied context condition. In spite of this advantage for the constant context condition during acquisition, the results showed a complete crossover in recall performance on the delayed retention test, with the varied context condition performing 22% better than the constant context condition. Forgetting, as measured by the drop in performance between the last retrieval practice trial and the final recall test, dropped 47%—a loss of 54% of the items recalled on the fifth retrieval practice trial. In contrast, forgetting in the varied context condition was negligible, a mere 3% drop, or 5% of the amount originally learned.

These findings clearly support the new theory of disuse. Retrieval practice was more difficult in the varied context condition, as evidenced by poorer acquisition performance in that condition relative to the constant context condition. More difficult retrieval practice resulted in better final recall, and almost no forgetting, even though the final recall test was given

after a 2-day retention interval and each face–name pair was tested in a new context. These results constitute evidence of decontextualization, because those who practiced with varied contexts were little affected by the context shifts at test, whereas those who always practiced with the original encoding context had highly contextualized learning that was undermined when those context cues were not provided at test.

Experiment 2

In Experiment 1 we found evidence consistent with our two predictions, that acquisition would benefit but retention would suffer from constant contexts during retrieval practice, relative to having retrieval practice in a new context every time. Although these results are consistent with the predictions derived from the new theory of disuse, an alternative explanation of the results of Experiment 1 might be that practicing retrieval with new contexts taught participants in the varied context condition to ignore the contexts, thereby decreasing the possibly distracting effects of new contexts. Thus, when the final test was given with new contexts for the face cues, varied context participants may have learned how to ignore the test context. Participants in the constant context condition, according to this alternative explanation, did not learn to ignore the practice context; therefore, they were distracted by the new contexts at test and showed poorer recall performance.

In Experiment 2, we tested these two alternative explanations of the results of Experiment 1 by testing final recall with no contexts present during the recall test. Thus, there were no new context videos to ignore on the retention test. According to the alternative artifact hypothesis, this change should eliminate the difference between the two treatment conditions on the final retention test in Experiment 2, because learning to ignore video contexts should confer no advantage when no video contexts are shown at test. The new theory of disuse hypothesis, however, predicts the same results that were observed in Experiment 1, namely, that the varied context group should outperform the constant context condition due to increased storage strength from retrieval practice that was done under conditions of low retrieval strength. In Experiment 2, as in Experiment 1, no acquisition contexts were provided on the 2-day delayed recall test; thus, retrieval strength on the retention test was assumed to be low, making that test a better index of encoding strength.

Method

Participants. Forty-eight Texas A&M University undergraduate students participated in this experiment in return for partial course credit. Participation was voluntary, and other options were available to earn equal credit. Participants self-enrolled in the experiment, and group assignment was based on time of testing. The number of participants in each experimental session depended upon the random number of participants who opted to enroll for a session, with a minimum of two and a maximum of 15 participants enrolling per session. There were 23 participants enrolled in the constant context condition and 25 participants in the varied context condition.

Design and materials. Background context during retrieval practice (constant or varied) was manipulated between subjects. The dependent measures were the proportion of names correctly recalled on each retrieval practice trial and the proportion of names correctly recalled on the final cued recall test 2 days later.

Procedure. The procedure for Experiment 2 was identical to the procedure described in Experiment 1, with one exception. For the final cued recall test, participants attempted to recall the correct name corresponding to each previously studied face. However, unlike in Experiment 1, where each test item (i.e., each face) was superimposed over a never-before-seen 5-s background scene, test items in Experiment 2 appeared over a blank white screen for 5 s.

Results

Acquisition. A 2×5 mixed ANOVA was computed, with context (constant and varied) as a between-subjects factor and retrieval practice trial (RP1, RP2, RP3, RP4, and RP5) as a repeated measure. The proportion of names correctly recalled on each retrieval practice trial served as the dependent measure. As in Experiment 1, there was a significant main effect of retrieval practice, $F(4, 184) = 309.51$, $MSE = 3.51$, $p < .0001$. There was also a significant main effect of context, $F(1, 46) = 13.79$, $MSE = 1.26$, $p < .001$; retrieval practice performance was greater for the constant condition than the varied condition (see Figure 2). Unlike in Experiment 1, the interaction between retrieval practice trial and background context was not significant, $F(4, 184) = 2.02$, $MSE = 0.02$, $p = .094$.

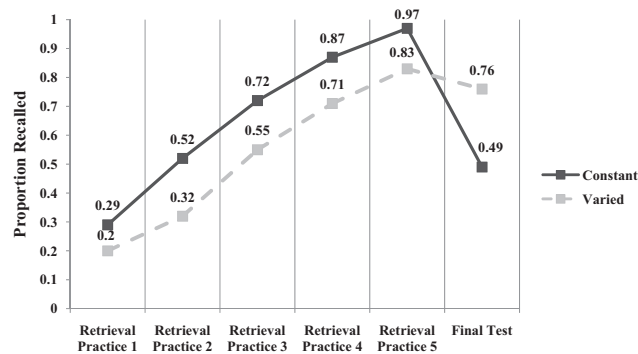


Figure 2. Proportion correctly recalled in Experiment 2 on each of five retrieval practice trials and on the final retention test as a function of context condition.

Retention and forgetting. A comparison of retention rates between the constant ($M = 0.49$, $SD = 0.27$) and varied ($M = 0.76$, $SD = 0.20$) context conditions revealed a significant difference in the proportion of face–name pairs correctly recalled on the final test, $t(46) = -3.98$, $p < .001$, Cohen's $d = 1.14$.

As in Experiment 1, forgetting was assessed in two ways. First, analyzing the absolute difference in the proportion of names correctly recalled in Retrieval Practice 5 and the final cued recall test revealed a significant difference, $t(46) = 7.60$, $p < .0001$, Cohen's $d = 2.29$. Participants in the constant context condition ($M = 0.47$, $SD = 0.25$) forgot a significantly greater proportion of face–name pairs than participants in the varied context condition ($M = 0.07$, $SD = 0.10$). Forgetting was also assessed in terms of the percentage of face–name pairs forgotten between Retrieval Practice 5 and the final cued recall test. This analysis also revealed significant differences between the constant ($M = 0.50$, $SD = 0.26$) and varied ($M = 0.09$, $SD = 0.12$) context conditions, $t(46) = 6.94$, $p < .0001$, Cohen's $d = 2.16$.

Discussion

The results of Experiment 2 replicate and extend the results of Experiment 1. Although there were no video contexts shown on the 2-day delayed recall test, Experiment 2, like Experiment 1, found that varying contexts during retrieval practice, relative to holding contexts constant across practice trials, impeded acquisition and benefited retention on the final recall test. Performance on retrieval practice trials was better in the constant context condition than the varied context condition, as in Experiment 1. On the last retrieval practice in Experiment 2, performance in the constant context condition was 15% better than that of the varied context condition. Again, there was a crossover in performance on the delayed retention test, with the varied context condition recalling 27% more names than participants the constant context condition. Performance in the constant context condition dropped 47% from the last retrieval practice trial on the 1st day to the final recall test 2 days later—a loss of half the items that had been successfully recalled on the last retrieval practice trial. There was barely any forgetting over the 2-day delay in the varied context condition, only a 5% drop, which was less than 10% of the originally learned names. Like the results of Experiment 1, these findings supported predictions derived from the new theory of disuse. The results also show that memories were decontextualized when face–name pairs were practiced in varied contexts, because those memories were barely affected when no context cues were provided on the retention test.

Experiment 3

Both Experiments 1 and 2 tested retention 2 days after retrieval practice trials, and both found greater retention for the varied context condition, relative to the constant context condition. The 2-day interval was used in those experiments so that memory could be tested under conditions that involved low retrieval strength. It is not clear how long a retention interval is needed to reduce retrieval strength at test sufficiently. In Experiment 3, we explored the same crossover interaction of constant and varied acquisition contexts with a much shorter retention interval than was used in Experiments 1 and 2. The procedure used in Experiment 3 was identical

to that described for Experiment 2, except that the final retention test was given only a few minutes after the acquisition phase, rather than 2 days later, as in Experiments 1 and 2. Using this brief retention interval, rather than the long one used in Experiments 1 and 2, we tested whether a 2-day retention interval, with context cues removed at test, would reduce retrieval strength enough to eliminate the varied and constant crossover interaction found in the first two experiments. No context videos were shown on the final retention test in Experiment 3.

Method

Participants. Twenty-one Texas A&M University undergraduate students participated in this experiment in return for partial course credit. Participation was voluntary, and other options were available to earn equal credit. Participants self-enrolled in the experiment, and group assignment was based on time of testing. The number of participants in each experimental session depended upon the random number of participants who opted to enroll for a session, with a minimum of two and a maximum of 15 participants enrolling per session. There were 12 participants in the constant context condition and nine in the varied context condition.

Design and materials. Context (constant or varied) was manipulated between subjects. The dependent measures included (a) the proportion of names correctly recalled at the end of the 1st day session and (b) the proportion of names correctly recalled on the final cued recall test 2 days later.

Procedure. The procedure for Experiment 3 was identical to that of Experiment 2, with one exception. Whereas Experiments 1 and 2 tested participants' memory for the face–name pairs after a 2-day retention interval, in Experiment 3 the final cued recall test was administered a few minutes after completing the fifth retrieval practice block. The retention interval corresponded to the amount of time taken to pass out materials for the final test and provide participants with instructions. As in Experiment 2, each test item was presented over a white background (i.e., no context).

Results

Acquisition. A 2×5 mixed ANOVA was computed, with context (constant and varied) as a between-subjects factor and retrieval practice trial (RP1, RP2, RP3, RP4, and RP5) as a repeated measure. The proportion of names correctly recalled on each retrieval practice trial served as the dependent measure. As in Experiments 1 and 2, there was a significant main effect of context, $F(1, 46) = 11.20$, $MSE = 1.47$, $p = .003$; retrieval practice performance was greater for the constant condition than the varied condition (see Figure 3). There was also a significant main effect of retrieval practice, $F(4, 76) = 94.47$, $MSE = 1.40$, $p < .0001$. Finally, the interaction between retrieval practice trial and background context approached significance, $F(4, 76) = 2.39$, $MSE = 0.04$, $p = .057$.

Retention and forgetting. An analysis of retention rates revealed significant differences in final test performance, $t(19) = -2.16$, $p = .043$, Cohen's $d = 0.96$. In line with Experiments 1 and 2, the varied context condition ($M = 0.73$, $SD = 0.18$) outperformed the constant context condition ($M = 0.54$, $SD = 0.21$) despite the fact that the retention interval between

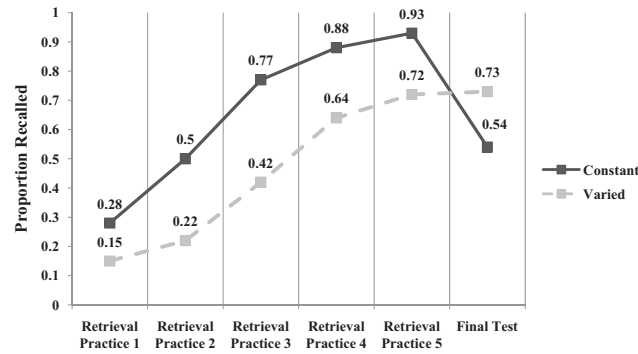


Figure 3. Proportion correctly recalled in Experiment 3 on each of five retrieval practice trials and on the final retention test as a function of context condition.

Retrieval Practice 5 and final test was reduced from 2 days to approximately 5 min.

As for forgetting, the constant ($M = 0.40$, $SD = 0.18$) and varied ($M = -0.01$, $SD = 0.10$) context conditions differed in terms of the proportion of face–name pairs forgotten between Retrieval Practice 5 and final test, $t(19) = 6.16$, $p < .0001$, Cohen’s $d = 2.93$. Similarly, when forgetting was analyzed in terms of the percentage of face–name pairs forgotten between Retrieval Practice 5 and final test, there were significant differences between the constant ($M = 0.43$, $SD = 0.20$) and varied ($M = -0.08$, $SD = 0.28$) context conditions, $t(19) = 4.88$, $p < .0001$, Cohen’s $d = 2.13$.

Discussion

The results of Experiment 3 replicate and extend the results of the two previous experiments. They show that constant context participants, who always practiced retrieval in the original encoding context, performed poorly when those context cues were not provided, even though the retention test was given only a few minutes later. In contrast, those in the varied context condition, who practiced retrieval of face–name pairs in new contexts on every trial, were not hindered by the removal of context cues on the final recall test. In the constant context condition, performance dropped 40% on the final recall test, with participants forgetting 43% of the face–name pairs that they had successfully recalled on the last retrieval practice trial. In the varied context condition, no forgetting was observed on the final recall test.

Why was forgetting so great in the constant context condition, even though the final retention test occurred only a few minutes after the final retrieval practice trial? No context cues were shown at test, so it may be that participants in the constant context condition were ignoring the face cues and attending instead to the context cues during acquisition. This explanation seems doubtful, however. As young adults, our participants have spent a lifetime learning the names of people, so the experimental task of learning face–name pairings cannot have been a new or novel task for them. The instructions to participants stated that they were to learn which name went with each face; no mention was made that contexts would or would not be present at test. Another explanation is that contextualization of learned material occurs because, in line with Hintzman’s (1986) ephoric model, rememberers might not clearly

distinguish to-be-learned materials from their contexts when contexts and to-be-learned items consistently co-occur, as in our procedure. In his oft-cited anecdote, for example, Mandler (1980) may not have differentiated the butcher’s face from his context; therefore, Mandler failed to recognize his butcher seen in a new context, that is, on a public bus.

Experiment 4

The results of Experiments 1–3 all show decontextualization and contextualization in learning, as defined by performance when supportive context cues are not provided at test. The results of all three experiments are consistent with predictions derived from the new theory of disuse. That is, they show that contextually cued retrieval practice, relative to varying context cues across practice trials, is not as effective for retention, if retention is tested with context cues. These results, however, are also consistent with predictions of the encoding variability theory, which states that varying contexts during acquisition results in more ways to remember a given item, thereby improving the odds that at least one of those varied encodings will be remembered on a final memory test (e.g., Gartman & Johnson, 1972; Melton, 1970; S. M. Smith, 1982). The constant context conditions in Experiments 1–3, by comparison, result in only a single way of remembering an item, and if the cue associated with that encoding is not provided at test, no other way of remembering that item remains. Both the new theory of disuse and the encoding variability theories can explain the decontextualization effects seen in the first three experiments of the present investigation.

Experiment 4 was conducted to test whether retrieval during practice was necessary to observe our decontextualization effects. The new theory of disuse depends on how retrieval practice can affect storage strength, with more difficult retrieval (due, for example, to a change of context cues) resulting in greater storage strength gains. In contrast, the encoding variability prediction holds whether acquisition trials involve retrieval practice or simply restudy of the critical items. The encoding variability explanation does not depend on retrieval practice; the same advantage gained by the varied context condition should occur, according to that theory, even if acquisition involves only restudy rather than retrieval practice. The new theory of disuse predicts that simply varying contexts on restudy trials (rather than retrieval practice

trials, as in Experiments 1–3) should have little effect on final retention, relative to a constant context restudy condition.

Method

Participants. Thirty-six Texas A&M University undergraduate students participated in this experiment in return for partial course credit. Participation was voluntary, and other options were available to earn equal credit. Participants self-enrolled in the experiment, and group assignment was based on time of testing. The number of participants in each experimental session depended upon the random number of participants who enrolled for a session, with a minimum of two and a maximum of 15 participants enrolling per session. There were 18 participants in the constant context condition and 18 in the varied context condition.

Design and materials. Background context (constant or varied) was manipulated between subjects. The dependent measures were (a) the proportion of faces correctly recalled on the five-item cued recall test at the end of the Day 1 session and (b) the proportion of faces correctly recalled on the final cued recall test on Day 2.

Procedure. The procedure for Experiment 4 was the same as that of Experiments 1–3, with two exceptions: First, participants did not perform retrieval practice after the initial study session. Instead, they restudied all 20 face–name pairs five times. The order of presentation was randomized in each of the five restudy blocks. As in Experiments 1–3, each face–name pair was superimposed over a 5-s movie scene. For participants in the constant context condition, these background videos corresponded to those they viewed during the initial study phase. For participants in the varied context condition, each face–name pair was coupled with a new background video scene for each of the five restudy blocks.

Second, to assess acquisition of the face–name pairs following the six study repetitions (i.e., initial study and the five restudy blocks), a four-item cued recall test was administered at the end of the Day 1 session. After a short delay in which participants were given test materials and instructions, participants were presented with four of the previously studied faces superimposed over new background contexts. The faces remained on the screen for 5 s, with a 2-s intertrial interval between test questions. Participants wrote down the correct name corresponding to each face, guessing if they were unsure. After the cued recall test, participants were excused for the day.

Participants returned 2 days later for a final cued recall test assessing their memory for all 20 face–name pairs presented with no background contexts (white background). The format of the cued recall test followed that of Experiments 2 and 3.

Results

Acquisition. Acquisition of the face–name pairs was assessed with an independent samples *t* test. There was no difference in cued recall between the constant ($M = 0.83$, $SD = 0.24$) and varied ($M = 0.71$, $SD = 0.31$) context conditions on the four-item cued recall test they participated in on an immediate test, $t(34) = 1.34$, $p = .189$ (see Table 1).

Retention and forgetting. Retention was examined in two ways. First, final test performance was assessed with data from the subset of four face–name pairs (previously tested on Day 1)

Table 1

Proportion Recalled in Experiment 4 on the Same-Day Test and 2-Day Delayed Test as a Function of Context Condition

| Test | Varied contexts | Constant contexts |
|-----------------|-----------------|-------------------|
| Same day | 0.71 | 0.83 |
| Two-day delayed | 0.36 | 0.51 |

included in the analysis. Although the difference in cued recall performance between the constant ($M = 0.51$, $SD = 0.32$) and varied ($M = 0.36$, $SD = 0.21$) context conditions was not significant, $t(34) = 1.72$, $p = .095$, there was a nonsignificant trend indicating an advantage for the constant context condition.

Final test performance was reanalyzed, excluding the subset of previously tested face–name pairs to control for potential retesting effects on those study items. This analysis yielded similar results. The constant ($M = 0.45$, $SD = 0.35$) and varied ($M = 0.30$, $SD = 0.21$) context conditions did not differ in final retention, $t(34) = 1.57$, $p = .126$.

To examine forgetting, a 2×2 mixed ANOVA assessed forgetting as a function of study context (constant vs. varied) and test (immediate vs. delayed), with the proportion of face–name pairs correctly recalled as the dependent measure. Critically, this analysis included final cued recall performance for the subset of four face–name pairs tested on Day 1. There was a main effect of test, $F(1, 34) = 62.96$, $MSE = 2.03$, $p < .0001$. Collapsed across context conditions, performance on the immediate cued recall test on Day 1 ($M = 0.77$, $SD = 0.28$) was superior to performance on the delayed test ($M = 0.44$, $SD = 0.27$). The main effect of study context was not significant, $F(1, 34) = 2.97$, $MSE = 0.35$, $p = .094$. Additionally, the interaction between test and context was not significant, $F(1, 34) = 0.11$, $MSE = 0.00$, $p = .745$.

The data were reanalyzed, excluding retested face–name pairs from final cued recall performance. The results of this 2×2 mixed ANOVA revealed a significant main effect of test, with immediate test performance ($M = 0.77$, $SD = 0.28$) once again superior to delayed test performance ($M = 0.38$, $SD = 0.29$), $F(1, 34) = 69.50$, $MSE = 2.80$, $p < .0001$. There was no effect of study context, $F(1, 34) = 2.84$, $MSE = 0.34$, $p = .101$. Finally, the interaction between test and study context was not significant, $F(1, 34) = 0.07$, $MSE = 0.00$, $p = .799$.

A final forgetting analysis directly compared memory performance for the subset of four face–name pairs tested both days as a function of study context condition. There was a main effect of test, $F(1, 34) = 8.67$, $MSE = 0.20$, $p = .006$. Immediate test performance ($M = 0.77$, $SD = 0.28$) was superior to final test performance ($M = 0.67$, $SD = 0.28$) for the subset of four face–name pairs. Neither the main effect of context, $F(1, 34) = 2.96$, $MSE = 0.38$, $p = .094$, nor the interaction between context and test, $F(1, 34) = 0.35$, $MSE = 0.01$, $p = .560$, reached statistical significance.

Discussion

The results of Experiment 4 showed no significant effects of varied and constant repetitions, either on an immediate recall test or on a 2-day delayed recall test. Although both immediate and delayed recall tests were given without context cues seen at study,

the advantage of the varied context condition over the constant context condition, which was observed in Experiments 1–3, was not evident in Experiment 4, which used multiple study trials, rather than multiple retrieval practice trials, as in the first three experiments. In fact, there was a slight but nonsignificant advantage in Experiment 4 for the constant context condition, contrary to the results of the first three experiments. Recall of items that had not been tested dropped nearly 40 percentage points from the initial test to the delayed test in both treatment conditions. These results are inconsistent with the encoding variability theory, which predicted a retention advantage for the varied context condition, relative to the constant context condition.

Could the failure to find a significant retention advantage for the varied context condition in Experiment 4 be due to a Type II statistical error that was caused by too little statistical power or an ineffective method? At least three facts argue against this interpretation of the results. First, it should be noted that in Experiment 4, the varied context condition did slightly worse than the constant context condition, not slightly better; more statistical power would not be expected to reverse that difference. Second, Experiment 4 used 36 participants; an a priori power analysis shows that to detect a varied context advantage of the same average magnitude seen in Experiments 1–3 (mean $d = 0.94$), 38 participants would be needed for an effect with $p = .05$ and statistical power at the recommended .80 level (see Cohen, 1988). Third, a nearly identical procedure produced clearly significant results in Experiments 1–3, which differed only in that retrieval practice trials were used during acquisition rather than restudy trials. Therefore, we are confident in concluding that the failure to find a varied context advantage in Experiment 4 was not due to a Type II error.

The results of Experiment 4 also add to the interpretation of the results of the previous experiments, because, as in those experiments, the retention test was given without the contexts seen during acquisition. In the first three experiments, performance in the constant context conditions dropped dramatically from the fifth retrieval practice to the final recall test, where no context cues were provided. In contrast, little or no forgetting was seen in the varied context conditions. Earlier, we speculated that these results might be explained by participants in the constant context condition attending to contexts rather than faces during acquisition, whereas those in the varied condition gave more selective attention at encoding to faces. This alternative explanation, however, should predict the same crossover interaction in Experiment 4. No interaction was observed in Experiment 4; forgetting was equally great in both treatment conditions.

Although it appears that our failure to find a varied context advantage was not due to a Type II statistical error, our failure to find a constant context advantage may have been due to such an error. Previous studies have shown a constant-encoding advantage (e.g., Bellezza & Young, 1989; Greene & Stillwell, 1995; Postman & Knecht, 1983). For example, Bellezza and Young (1989) presented pairs of words twice, that is, with the same response word in both pairs. The stimulus words, which served as verbal contexts for response terms, were either the same on both presentations (e.g., *bottle–tower*, *bottle–tower*) or different on the two presentations (e.g., *hotel–stone*, *ticket–stone*). Cued recall was better in the constant encoding condition (i.e., the same stimulus term twice) than in the varied condition (varied stimulus terms) when repetitions were spaced. Greater statistical power (i.e., more par-

ticipants) in Experiment 4 of the present study might have found a similar advantage in the constant context condition. Consistent with this possibility, a post hoc power analysis shows that our failure to reject the hypothesis that participants in the constant context condition recalled no more than those in the varied context condition may be due to insufficient power. With 36 participants in Experiment 4, our observed power was .48. To attain the recommended power of .80, a sample size of 84 participants would be needed. It should be noted that, although Experiment 4 has insufficient power to detect a constant context advantage, it nonetheless had sufficient power to detect a varied context advantage of the same magnitude as the effects observed in Experiments 1–3.

General Discussion

Four experiments were reported that examined decontextualization of memories, that is, memory of newly learned material that does not rely on the reinstatement of a particular context. In Experiments 1–3, names associated with unfamiliar faces were recalled more poorly during retrieval practice trials when they were tested in a new context on every trial, the varied context condition, relative to testing retrieval in the original encoding context on every trial, the constant context condition. When tested in new contexts, or with no contexts at test, however, names learned in the varied context condition were recalled better than names acquired in the constant context condition. Although forgetting was great following constant context acquisition, there was virtually no forgetting for names learned in varied contexts, even when testing occurred 2 days later, when encoding contexts were absent at test, and when testing occurred with new, unfamiliar contexts (see Table 2). These results show that material practiced in constant, unchanging contexts remain contextualized; that is, those memories are vulnerable to changes in contexts. In comparison, recall of material practiced in varied contexts was more decontextualized, that is, resilient even when contexts changed, showing good recall and little forgetting.

The results were predicted by, and are consistent with the new theory of disuse, which states that encoding strength is increased more by difficult retrieval practice, relative to easier retrieval practice. Recall of paired associates tested in new contexts that were different from encoding contexts was shown to be worse than recall tested with appropriate encoding contexts reinstated (S. M. Smith et al., 2014). Therefore, retrieval practice in a new context,

Table 2
Forgetting, Calculated as Percentage Drop From Retrieval Practice 5 to the Final Recall Test, for Experiments 1–4

| Experiment | Treatment | Condition | Forgetting (%) |
|------------|------------------------------------|-----------|----------------|
| 1 | Two-day delayed test, new contexts | Constant | 47 |
| | | Varied | 3 |
| 2 | Two-day delayed test, no contexts | Constant | 47 |
| | | Varied | 7 |
| 3 | Same-day test, no contexts | Constant | 39 |
| | | Varied | –1 |
| 4 | Six study, no retrieval practice | Constant | 32 |
| | | Varied | 35 |

as in the varied condition, should be more difficult than retrieval practice in the original encoding context.

Although the results of all four experiments are consistent with the new theory of disuse, they are not all consistent with the theory of encoding variability. Experiment 4 was conducted as a critical test of these two theories. Predictions drawn from the encoding variability theory do not rely on the distinction between retrieval practice trials and restudy trials at acquisition; that is, the advantage of varied contexts during acquisition should occur regardless of the manner of reexposure to the material being learned. The new theory of disuse, however, predicts that retrieval practice has a greater benefit for encoding strength when retrieval is more difficult, as in the varied context condition. No clear prediction can be drawn from the new theory of disuse with respect to varied and constant contexts if retrieval practice is not involved during acquisition, as in Experiment 4. Therefore, the finding that varied contexts, relative to constant contexts, led to better retention and less forgetting only when retrieval practice was involved during acquisition (Experiments 1–3), and not when restudy trials were used (Experiment 4), is consistent with predictions of the new theory of disuse and not the theory of encoding variability. It is not clear how to resolve the results of Experiment 4 with S. M. Smith et al.'s (1978) finding that a list of paired associates repeated in two environmental contexts were recalled better, relative to two presentations in the same context. There are many differences between Experiment 4 of the present investigation and Experiment 1 in S. M. Smith et al.'s study, including differences in the types of contexts used, the ratio of items to contexts in the two studies, the number of repetitions during acquisition, and the intervals between study trials. It is also the case that not all studies of the varied context superiority effect (S. M. Smith et al., 1978) have found the effect. Although Isarida and Isarida (2010) replicated the effect when simple contexts were used, they found a constant context advantage when more complex contexts were used (Isarida & Isarida, 2005, 2010). Further research will be necessary to resolve the differences among these studies, and between study and retrieval practice effects in varied and constant contexts.

There are several alternative explanations, other than the new theory of disuse, that might be used to explain the present findings. One possibility is that retrieval practice in varied contexts might simply create stronger, more robust memories, rather than memories that are decontextualized. For example, studying face–name pairs in new contexts might confer greater distinctiveness to those items, making them more memorable. A similar hypothesis is that varied contexts add arousal to study episodes, thereby making items learned under conditions of increased arousal more memorable. Two aspects of the present results, however, cast doubt on this type of interpretation. One problem with this “stronger memories” hypothesis is that it would seem to predict that retrieval practice in varied contexts, relative to constant contexts, would produce better performance during acquisition, the reverse of what was observed in Experiments 1–3. Another problem with such an interpretation is that it would predict superior memory for the varied context condition in Experiment 4, which also was not observed.

Another explanation of our results could be a transfer-appropriate processing account (e.g., Blaxton, 1989; Morris, Bransford, & Franks, 1977). Transfer-appropriate processing is the idea that performance on a memory test depends on the degree to

which the cognitive processes required by the test correspond to the processes used during study. In terms of the present experiments, if material is practiced in a new context on each study trial, as in the varied context condition, then a test conducted with new contexts present, as in Experiment 1, should produce optimal performance, according to the transfer-appropriate processing theory. For material practiced in a single (constant) context, testing in a new context, according to this explanation, should produce relatively poor memory. These were exactly the results observed in Experiment 1. The results of Experiments 2 and 3, however, in which no contexts were presented at test, are not consistent with a transfer-appropriate processing account, because video contexts were presented during acquisition, but not at test. The results of Experiment 4, in which the constant context condition produced somewhat better (and clearly not worse) memory than the varied context condition, also are not easily explained by the transfer-appropriate processing account.

A hypothesis that is not rejected by the present results derives from Benjamin's theory of reminding (e.g., Benjamin & Tullis, 2010; Benjamin & Ross, 2011). This theory states that conditions that make current and past occurrences of an item more dissimilar (such as varied contexts, relative to constant contexts) lead to greater memory enhancement if the current repeated event successfully reminds the learner of the previously studied item. Strengthening this explanation of our results is the fact that in Experiments 1–3, corrective feedback was given immediately following each retrieval practice, helping to ensure reminders of earlier presentations of repeated items. Furthermore, no such reminders occurred in Experiment 4; restudy occurrences in varied contexts may have impeded reminders of previous study trials, and in that experiment varied contexts did not benefit memory. Thus, although the present experiments were not motivated by the reminding theory, the results of all four experiments appear to be consistent with that theory.

Another explanation of the present results is that participants in the constant context conditions of the present experiments may have been using the background contexts as a “contextual crutch” in the sense that they may have relied on contextual associations, rather than the faces, to help them recall the names that were paired with the faces. On the one hand, there may be some good reasons to reject this contextual crutch hypothesis. Adults, including the student participants in the present study, spend most of their lives learning the names of individuals, such as friends, acquaintances, and celebrities. Although participants may not be experts at learning names of individuals (most may find it quite difficult to do so), it is reasonable to assume that they understand what it means to learn a person's name, and it is a task used in other studies of the effects of retrieval practice on memory (e.g., Carpenter & DeLosh, 2006). Instructions in our experiments made it clear that the participants' task was to remember the name that goes with each face. Furthermore, the video contexts were not obviously related to the faces in any way, so contexts would not be expected to support new associations the way that, for example, semantically associated materials might. On the other hand, our context videos were attention-getting, and participants in the constant context conditions may have paid more attention to the videos than may have been warranted for successful performance on the task. Perhaps the effect of testing retrieval in varied contexts in our experiments was to teach participants to ignore the contexts during the learning

phase, helping participants better focus their learning efforts on the face–name pairs. Although the use of final retention tests that had no context cues (Experiments 2 and 3) ruled out the possibility that varied practice contexts protected participants from being distracted at test, it may be, nonetheless, that varied-context participants were protected from potentially distracting contexts during retrieval practice trials. It could also be said that in naturalistic situations, such as learning the name of one’s butcher, who is always seen in the same context, a meat shop environment, people also use the constant environment as a contextual crutch associated with the learned name. As such, our constant context conditions may well simulate the contextualized learning that explains something similar to Mandler’s (1980) butcher-on-the-bus experience.

The effects at test of withdrawing or replacing study contexts in the present experiments are noteworthy. The context-dependent memory effects seen on the first retrieval practice trial in Experiments 1–3 of the present study are consistent with the robust effects reported by S. M. Smith and Manzano (2010) and S. M. Smith et al. (2014), studies that also used video recordings of environments as contexts. That is, in the three experiments of the present investigation, the original study context was either reinstated (in the constant context conditions) or changed (in the varied context conditions) on the first retrieval practice trial, eliciting reinstated context advantages. Especially noteworthy is that retention for material practiced five times in constant context conditions showed sizable forgetting effects when study contexts were changed or removed at test, even in Experiment 3, when the retention test was given only a few minutes after the fifth retrieval practice. These results demonstrate not only compelling effects of video contexts but also a clear limitation of the beneficial effects of retrieval practice on memory (e.g., Roediger & Karpicke, 2006).

There may be some applications of the methods used in the present investigation, particularly in the area of learning new vocabulary terms, such as an unfamiliar language or subject domain. In the present experiments there were no obvious semantic relations between contexts and learned material. It might be useful to support initial learning with semantically related contexts. For example, learning that a word in an unfamiliar language means *pilot* might be augmented, at least in early learning, if practice trials occur with an aviation context displayed in the background. Such semantic support might help initial learning, but the results of the present experiments suggest that subsequent variation of contexts would be important for the new learning to be decontextualized, that is, retained without contextual support.

References

- Bellezza, F. S., & Young, D. R. (1989). Chunking of repeated events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 990–997. doi:10.1037/0278-7393.15.5.990
- Benjamin, A. S., & Ross, B. H. (2011). The causes and consequences of reminding. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 71–105). New York, NY: Psychology Press.
- Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*, 228–247. doi:10.1016/j.cogpsych.2010.05.004
- 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.
- Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 657–668. doi:10.1037/0278-7393.15.4.657
- Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, *34*, 268–276. doi:10.3758/BF03193405
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Craik, F. I. M. (1979). Levels of processing: Overview and closing comments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 447–461). Hillsdale, NJ: Erlbaum.
- Farnsworth, P. R. (1934). Examinations in familiar and unfamiliar surroundings. *Journal of Social Psychology*, *5*, 128–129. doi:10.1080/00224545.1934.9921593
- Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, *13*, 333–345. doi:10.3758/BF03202501
- Finley, J. R., Benjamin, A. S., Hays, M. J., Bjork, R. A., & Kornell, N. (2011). Benefits of accumulating versus diminishing cues in recall. *Journal of Memory and Language*, *64*, 289–298. doi:10.1016/j.jml.2011.01.006
- Gartman, L. M., & Johnson, N. F. (1972). Massed versus distributed repetition of homographs: A test of the differential encoding hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *11*, 801–808. doi:10.1016/S0022-5371(72)80016-1
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*, 325–331. doi:10.1111/j.2044-8295.1975.tb01468.x
- Greene, R. L., & Stillwell, A. M. (1995). Effects of encoding variability and spacing on frequency discrimination. *Journal of Memory and Language*, *34*, 468–476. doi:10.1006/jmla.1995.1021
- Gruppuso, V., Lindsay, S., & Masson, M. E. J. (2007). I’d know that face anywhere! *Psychonomic Bulletin & Review*, *14*, 1085–1089. doi:10.3758/BF03193095
- Hayes, S. M., Nadel, L., & Ryan, L. (2007). The effect of scene context on episodic object recognition: Parahippocampal cortex mediates memory encoding and retrieval success. *Hippocampus*, *17*, 873–889. doi:10.1002/hipo.20319
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review*, *93*, 411–428. doi:10.1037/0033-295X.93.4.411
- Hockley, W. E. (2008). The effects of environmental context on recognition memory and claims of remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1412–1429. doi:10.1037/a0013016
- Isarida, T., & Isarida, T. K. (2005). Study-time effect on free recall within and out of context. *Memory*, *13*, 785–795. doi:10.1080/09658210500139218
- Isarida, T., & Isarida, T. K. (2010). Effects of simple- and complex-place contexts in the multiple-context paradigm. *Quarterly Journal of Experimental Psychology*, *63*, 2399–2412. doi:10.1080/17470211003736756
- Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review*, *120*, 852–872. doi:10.1037/a0034246
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271. doi:10.1037/0033-295X.87.3.252
- Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*, 596–606. doi:10.1016/S0022-5371(70)80107-4

- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. doi:10.1016/S0022-5371(77)80016-9
- Pan, S. (1926). The influence of context upon learning and recall. *Journal of Experimental Psychology*, *9*, 468–491. doi:10.1037/h0073472
- Postman, L., & Knecht, K. (1983). Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior*, *22*, 133–152. doi:10.1016/S0022-5371(83)90101-9
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255. doi:10.1111/j.1467-9280.2006.01693.x
- Saufley, W. H., Jr., Otaka, S. R., & Bavaresco, J. L. (1985). Context effects: Classroom tests and contextual independence. *Memory & Cognition*, *13*, 522–528. doi:10.3758/BF03198323
- Schacter, D. L., Eich, J. E., & Tulving, E. (1978). Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *17*, 721–743. doi:10.1016/S0022-5371(78)90443-7
- Semon, R. (1923). *Mnemic psychology* (B. Duffy, Trans.). London, England: Allen & Unwin. (Original work published 1909)
- Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 460–471. doi:10.1037/0278-7393.5.5.460
- Smith, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory & Cognition*, *10*, 405–412. doi:10.3758/BF03197642
- Smith, S. M. (1994). Theoretical principles of context-dependent memory. In P. Morris & M. Gruneberg (Eds.), *Theoretical aspects of memory* (2nd ed., pp. 168–195). New York, NY: Routledge.
- Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*, 342–353. doi:10.3758/BF03197465
- Smith, S. M., Handy, J. D., Angello, G., & Manzano, I. (2014). Effects of similarity on environmental context cueing. *Memory*, *22*, 493–508. doi:10.1080/09658211.2013.800553
- Smith, S. M., & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods*, *42*, 292–301; doi:10.3758/BRM.42.1.292
- Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and a meta-analysis. *Psychonomic Bulletin & Review*, *8*, 203–220. doi:10.3758/BF03196157
- Smith, S., & Guthrie, E. R. (1921). *General psychology in terms of behavior*. New York, NY: Appleton.
- Staudigl, T., & Hanslmayr, S. (2013). Theta oscillations at encoding mediate the context-dependent nature of human episodic memory. *Current Biology*, *23*, 1101–1106. doi:10.1016/j.cub.2013.04.074
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 382–402). New York, NY: Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.

Received December 4, 2013

Revision received March 18, 2014

Accepted March 19, 2014 ■