The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Explaining retrieval-induced forgetting: A change in mental context between the study and restudy practice phases is not sufficient to cause forgetting

Dorothy R. Buchli, Benjamin C. Storm & Robert A. Bjork

To cite this article: Dorothy R. Buchli, Benjamin C. Storm & Robert A. Bjork (2016) Explaining retrieval-induced forgetting: A change in mental context between the study and restudy practice phases is not sufficient to cause forgetting, The Quarterly Journal of Experimental Psychology, 69:6, 1197-1209, DOI: 10.1080/17470218.2015.1076866

To link to this article: http://dx.doi.org/10.1080/17470218.2015.1076866

Accepted author version posted online: 04 Aug 2015.
Published online: 26 Aug 2015.

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2016
Vol. 69, No. 6, 1197–1209, http://dx.doi.org/10.1080/17470218.2015.1076866

Routledge
Taylor & Francis Group

Explaining retrieval-induced forgetting: A change in mental context between the study and restudy practice phases is not sufficient to cause forgetting

Dorothy R. Buchli[1], Benjamin C. Storm[2], and Robert A. Bjork[3]

[1]Department of Psychology, Mercer University, Macon, Georgia
[2]Department of Psychology, University of California, Santa Cruz, USA
[3]Department of Psychology, University of California, Los Angeles, USA

(Received 2 February 2015; accepted 28 June 2015; first published online 27 August 2015)

Retrieving information can impair the subsequent recall of related information. Such retrieval-induced forgetting is often attributed to inhibitory mechanisms, but Jonker, MacLeod, and Seli (2013) recently proposed an alternative account. In their view, the study and retrieval-practice phases constitute two disparate contexts, and impairment of unpractised members from practised categories is attributable to their being absent from the retrieval-practice context, which is where, according to Jonker et al., participants preferentially search at the time of final test. In evidence of this account, Jonker et al. showed that even restudy practice—which is assumed by the inhibitory account to be insufficient to cause forgetting (i.e., retrieval-specificity)—can cause forgetting when a mental context change is inserted between study and restudy. The present research sought to replicate this finding while also testing the possibility that a far mental context change would cause more forgetting than a near mental context change. In Experiment 1, participants described a vacation inside the United States (near) or outside the United States (far). In Experiments 2 and 3, participants described the layout of their own home (near) or their parents' home (far). In contrast to the predictions of the context account, however, but consistent with the predictions of the inhibitory account, none of the restudy-plus-context-change conditions resulted in significant forgetting.

Keywords: Retrieval-induced forgetting; Contextual-cuing account; Inhibition; Context change.

Forgetting, though often regarded as a frustrating or maladaptive failure of memory, is an adaptive process that is essential for successful remembering. The task of recalling relevant information that is pertinent in the present would be difficult or impossible, for example, without some way of setting aside outdated or irrelevant information. One mechanism that has been proposed to underlie such adaptive forgetting is inhibition (e.g., Anderson, 2003; Bjork, 1989). The basic idea is that retrieving some target information from memory requires not only selecting that information, but also selecting against competing information—that is, other information associated with

Correspondence should be addressed to Dorothy R. Buchli, Department of Psychology, Mercer University, 1400 Coleman Avenue, Macon, GA 31207, USA. E-mail: buchli_dr@mercer.edu

The authors offer sincere thanks to Tanya Jonker and Collin MacLeod for their insightful comments and continued support and assistance in preparing this manuscript for publication. They are similarly gratefully indebted to Michael Anderson and Jonathan Fawcett for their encouragement, valuable feedback, and helpful guidance.

the same cue or cues, which presumably becomes activated and competes for access. This inhibition has been argued to explain a rather unintuitive empirical observation—that retrieving some items from memory causes the forgetting of related items, a phenomenon known as retrieval-induced forgetting (Anderson, Bjork, & Bjork, 1994).

Retrieval-induced forgetting is a fairly robust phenomenon. It has been observed with a variety of materials and in a number of applied situations (for an extensive review of retrieval-induced forgetting in various applied settings, see Storm et al., 2015). Studies of retrieval-induced forgetting typically employ a three-phase retrieval-practice paradigm. During study, participants are presented with a series of category–exemplar pairs drawn from several categories (e.g., *fruits–orange, drinks–rum, professions–nurse*). Subsequently, during retrieval practice, participants are asked to repeatedly retrieve half of the items from half of the categories in response to selective retrieval cues (e.g., *fruits: or___, drinks: ru___*). After a brief delay, participants are tested on all of the items, often via category-plus-one-letter-stem retrieval cues (e.g., *fruits: o___, drinks: r___*). The retrieval-practice paradigm produces three types of items: *Rp+* items refer to practised items from practised categories; *Rp−* items refer to unpractised items from practised categories; Nrp items refer to items from nonpractised categories.[1]

Not surprisingly, Rp+ items are recalled better at test than Nrp items, a finding that is consistent with research on the testing effect (see, e.g., Bjork, 1975; Roediger & Karpicke, 2006). Interestingly, recall for Rp− items is impaired relative to Nrp items. This decrement in recall for Rp− items relative to Nrp items reflects the phenomenon of retrieval-induced forgetting. Critically, the term retrieval-induced forgetting refers to an empirical effect (that recalling a subset of information impairs subsequent recall for related information)—it does not stipulate the mechanism presumed to underlie the effect.

The theoretical explanations that have been put forth to explain retrieval-induced forgetting can be broadly grouped into *inhibition-based theories* and *competition-based theories*. For example, whereas inhibition-based theories assume that an active control mechanism is recruited during retrieval practice to suppress the accessibility of competing information in order to facilitate the retrieval of target information, and that it is this inhibition that renders Rp− items less recallable than Nrp items (Anderson, 2003; Anderson et al., 1994; Storm & Levy, 2012), competition-based theories assume that retrieval-induced forgetting can be explained by strength-based competition at test and other non-inhibitory mechanisms (Raaijmakers & Jakab, 2013; Verde, 2012). For example, the retrieval of a subset of items may strengthen those items and cause them to interfere with, or block the recall of, weaker items, thus preventing them from becoming accessible at test. Although strength-based interference does probably play some role in observations of retrieval-induced forgetting, there is now substantial evidence implicating a role for inhibition as well (for recent qualitative and quantitative reviews, see Storm & Levy, 2012; Murayama, Miyatsu, Buchli, & Storm, 2014, respectively).

Recently, an intriguing new account of retrieval-induced forgetting has been put forth. This account, referred to as the context-based account (Jonker, Seli, & MacLeod, 2013), contends that retrieval-induced forgetting is a consequence of inappropriate contextual cuing at test. The study and retrieval-practice phases in the retrieval-practice paradigm are assumed to represent two disparate contexts. Because Nrp categories are only encountered in the study phase, Nrp items are only associated with the study context. In contrast, Rp categories are encountered in both the study phase and the retrieval-practice phase. Thus, when cued with a retrieval practice category cue at final test, participants may inappropriately search for Rp− items in the retrieval-practice context

[1]Post-experimental questionnaires were administered asking participants where they lived and with whom. Participants who still lived with their parents were removed from the analysis. Furthermore, the same pattern of results emerged when those participants were included.

while not doing so for Nrp items, making it relatively more difficult to access Rp− items than Nrp items and thus leading to retrieval-induced forgetting. Said differently, because the categories associated with Rp− items were practised during retrieval practice, participants may search the retrieval-practice context for Rp− items at test, rendering those items less accessible not because they were inhibited, but because participants were not able or inclined to effectively target the appropriate context in their attempt to recall them. Nrp items, on the other hand, would not suffer this type of inappropriate contextual cuing because items in the Nrp category were only encountered in the study context, making it more likely that subjects would reinstate the study context when attempting to recall those items.

Jonker et al. (2013) conducted several experiments to garner evidence in support of the context-based account. In one experiment, participants were asked to study a series of category–exemplar pairs in the initial study phase and then to restudy a subset of those pairs prior to final test. Although this sort of extra-study practice has typically not been shown to cause the forgetting of related non-practised items (Anderson, Bjork, & Bjork, 2000; Bäuml, 2002; Jonker et al., Experiment 1)—a result often cited as evidence against competition-based accounts and in support of inhibitory-based accounts (Anderson, 2003; Storm & Levy, 2012)—Jonker et al. argued that the lack of forgetting might be attributed to the fact that restudy practice typically fails to induce a change in context between study and practice. That is, the initial study phase and the restudy phase may be represented as one large context, presumably because retrieval practice induces a shift in context owing to the change in task demands and processing between study and retrieval, whereas restudy does not. Jonker et al. predicted that by implementing a context shift between study and restudy—leading, in their view, to the study and restudy phases becoming represented as two separate contexts—non-practised items associated with restudied categories would suffer forgetting, which is exactly what the authors found. Specifically, by inserting a mental context change manipulation between study and restudy (i.e., asking participants to imagine their parents' house and draw a diagram of the layout), restudy practice caused non-practised items from practised categories to suffer significant forgetting.

In a follow-up experiment (Experiment 2b), Jonker et al. (2013) examined whether context reinstatement might eliminate the effect of forgetting caused by extra study. Specifically, the authors inserted a reinstatement task immediately prior to the final test. In this task, participants were asked a series of questions that were designed to encourage them to think about the study phase at the beginning of the experiment. For example, "What did you notice when you first entered the room for the experiment?" They also employed the Star Wars theme song at the beginning of the experiment to provide a kind of distinct anchor to signal the beginning of the experiment. As predicted, when participants engaged in this reinstatement task, no retrieval-induced forgetting was observed, presumably because participants were now able to effectively target the study phase at test when attempting to retrieve Rp− items, thus reducing the costs associated with inappropriate contextual cuing. In the same paper, Jonker et al. report a successful replication of both the initial experiment showing a significant RIF-like effect following restudy-plus context change and the finding that this RIF-like effect is abolished following reinstatement of the study context.

In their third experiment, Jonker et al. (2013) employed the standard retrieval-practice paradigm in which participants first studied category–exemplar pairs and then performed retrieval practice on half of the items from half of the categories. However, during study and retrieval practice, the items were presented along with videos depicting everyday contexts, such as the first-person perspective of walking downstairs or a panoramic view of a kitchen. These videos have been demonstrated to reliably induce environmental context effects (Smith & Manzano, 2010). The pairs were first studied along with one context video. That is, all items in one category were paired with the same video. Then, during retrieval practice, the practised cues were paired with a new video. At test, one of

the two videos was provided. Presumably, when participants received the retrieval-practice video, they would be likely to search the retrieval-practice context. When participants received the study video, however, they would be likely to search the study phase, thus reinstating the study context and reducing the effects of inappropriate contextual cuing. In support of this assumption, and of the context-based account, retrieval-induced forgetting was observed in the former condition where the retrieval-practice context was reinstated, but not in the latter condition where the study context was reinstated.

Logic of the present studies

The findings of Jonker et al. (2013) suggest that context, and particularly the contextual cues participants sample at test, play a critical role in determining the occurrence of restudy-induced forgetting and retrieval-induced forgetting. In the present research, we sought to replicate and extend one of the critical findings observed by Jonker et al.—specifically, that inducing a change in context between study and extra study practice is sufficient to cause non-practised items from practised categories to be forgotten. In addition to replicating this finding, we sought to extend it by examining whether the magnitude of the forgetting effect would be influenced by the magnitude of the contextual shift between study and restudy practice.

To investigate this possibility, we borrowed two context manipulations employed by Delaney, Sahakyan, Kelley, and Zimmerman (2010). In one manipulation—which we adapted for Experiment 1—participants were asked to imagine either visiting their family within their home country (near-imagination task) or going on a vacation outside their home country (far-imagination task). In the other manipulation—which we adapted for Experiments 2 and 3, and which was very similar to the context manipulation used by Jonker and colleagues—participants were asked to either imagine the layout of their own home (near-imagination) or of their parents' home (far-imagination). Delaney et al. found that the far-imagination tasks disrupted memory performance for a previously studied word list to a greater extent than the near-imagination task, presumably because the far-imagination tasks led to a stronger shift in context and thus a reduced ability of participants to reinstate the original study context at test. Based on these results, in the present context, one might expect that separating the study and restudy practice phases with a far-imagination task would lead to a larger forgetting effect than separating them with a near-imagination task. On the other hand, any form of context shift may be sufficient for non-practised items from practised categories to suffer forgetting, in which case the near and far imagination tasks may be equally effective in causing forgetting.

EXPERIMENTS 1 AND 2

In summary, the aims of the first two experiments were twofold: (1) to again replicate the finding observed by Jonker et al. (2013) that inserting a context change manipulation between study and restudy practice results in the forgetting of non-practised items from practised categories, and (2) to determine if the magnitude of the context change manipulation dictates the extent to which such forgetting is observed. To investigate these issues, four between-subject conditions were employed in both Experiments 1 and 2: a typical retrieval-practice group, a restudy group without context change, a restudy group with near-imagination context change, and a restudy group with far-imagination context change. Based on Jonker et al.'s context account, the restudy groups without context change should fail to exhibit forgetting, whereas the retrieval-practice groups and the restudy groups with context change should exhibit significant forgetting. Moreover, based on the results of Delaney et al. (2010), we endeavoured to see if the magnitude of the context shift determines the degree to which participants become prompted to search the inappropriate context at test. If it does, then participants in the far-context-change groups should exhibit greater levels of forgetting than participants in the near-context-change groups.

Method

Participants and design

In total, 480 (240 in Experiment 1 and 240 in Experiment 2) students from the University of California, Los Angeles (UCLA) participated for credit in an introductory psychology course. Retrieval-Practice Status (Nrp vs. Rp− vs. Rp+) was manipulated within subjects. Experimental condition (no-context-change restudy vs. no-context-change retrieval practice vs. near-context-change restudy vs. far-context-change restudy) was manipulated between subjects.

Materials

Eight categories were selected, each consisting of six high-frequency exemplars, for a total of 48 category–exemplar pairs (taken directly from Anderson et al., 1994). The pairs were counterbalanced such that each item served equally often as an Rp+ item, Rp− item, and Nrp item.

Procedure

Participants were randomly assigned to one of the four conditions described below.

No-context-change restudy. This condition was identical for participants in Experiments 1 and 2. During the initial study phase, 48 exemplar pairs were presented via computer at a rate of one pair every 4 s. Order was set randomly, with the constraint that no two consecutive pairs could be shown from the same category. Immediately following the completion of the study phase, participants were prompted to restudy half of the exemplars from half of the categories. There were three rounds of practice such that the 12 restudied pairs were presented three times each for 7 s, resulting in a total of 36 restudy trials. After the restudy phase, there was a 5-min retention interval, during which participants attempted to complete a series of mathematics problems. A category-plus-one-letter-stem cued recall test was then administered, in which participants were given 6 seconds to recall each exemplar. To control for output interference, the final test was divided into two test blocks, with Rp− items and half of the Nrp items (Nrp− items) tested in the first block, and Rp+ items and the other half of the Nrp items (Nrp+ items) tested in the second block. The particular set of Nrp items serving as Nrp− and Nrp+ items was counterbalanced across participants.

Near-context-change restudy. This condition was the same as the no-context-change restudy condition except for one important difference. Specifically, in Experiment 1, between the study and restudy phases, participants were asked to describe a vacation within the United States that had taken place within the past three years. A recruitment procedure was implemented, such that participants were selected only if they had taken a vacation outside of their home state in the past 3 years. Participants were prompted to describe what they saw, felt, smelled, and experienced with all of their senses during the vacation. In Experiment 2, participants were asked to write a detailed description of the interior and exterior of their home as they mentally walked through each of the rooms. They were given one minute to complete this task. All participants completed the experiment in a single room, accompanied by an experimenter to ensure that they complied appropriately with all instructions.

Far-context-change restudy. This condition was also the same as the no-context-change restudy condition, except for what participants did between the study phase and the restudy phase. In Experiment 1, participants were asked to describe a vacation outside of the United States had taken took place within the past 3 years. Once again, only individuals who met this requirement were allowed to participate. International exchange students were also excluded. In Experiment 2, participants were asked to provide a detailed description of the exterior and interior of their *parents*' home as they mentally walked through each room. All participants were accompanied by an experimenter to ensure that they complied with instructions.

No-context-change retrieval practice. This condition, which was the same in Experiments 1 and 2,

was identical to the no-context-change restudy condition, except that participants were given retrieval practice for half of the exemplars from half of the categories instead of restudy practice. Specifically, as typically employed in studies of retrieval-induced forgetting, participants were given category-plus-two-letter stem cues for 7 s each (e.g., fruit-or____) and asked to recall the appropriate exemplar from the appropriate studied category. As in the restudy conditions, there were three blocks of practice, with participants attempting to retrieve each of the 12 to-be-practised exemplars three times each, resulting in a total of 36 trials.

For the sake of simplicity, and to facilitate the readers' comprehension of our results, we refer to practised and non-practised items from practised categories as Rp+ and Rp− items, respectively, regardless of the nature of the practice that participants performed.

Results

Retrieval-practice performance

Participants recalled the appropriate exemplar on 89% and 90% of the retrieval-practice trials in Experiment 1 and 2, respectively.

Final recall performance for practised items and baseline controls

As can be seen in Figure 1, a significant facilitation effect was observed in all four conditions in Experiment 1 and 2, such that Rp+ items were recalled significantly better than were Nrp+ items (all $p < .001$).

Final recall performance for non-practised items from practised categories and baseline controls

Recall performance in Experiments 1 and 2 for Rp− and Nrp− items on the final test is shown as a function of condition in Figure 1 and was analysed using a 2 (Rp− vs. Nrp−) × 4 (Near vs. Far vs. Retrieval Practice vs. Restudy) mixed Analysis of Variance (ANOVA). In Experiment 1, although the main effects of retrieval-practice status, $F(1, 236) = .00$, $MSE = .02$, $p = .97$, and context-change condition, $F(1, 236) = .03$, $MSE = .05$, $p = .61$, were not significant, a significant interaction was observed, $F(3, 236) = 2.71$, $MSE = .02$, $p < .05$. In Experiment 2, the effects of retrieval-practice status, $F(1, 236) = 19.01$, $MSE = .31$, $p = .01$, and context change condition, $F(1, 236) = 4.18$, $MSE = .18$, $p = .01$, were significant, as was the interaction, $F(3, 236) = 63.98$, $MSE = 1.06$, $p = .01$.

As can be seen in Figure 1, in Experiment 1, forgetting was observed in the retrieval-practice condition, with Rp− items ($M = .66$, $SE = .02$) recalled significantly less well than Nrp items ($M = .72$, $SE = .02$) items, $t(59) = 2.00$, $p < .05$, $d = .36$. The same pattern was observed in Experiment 2 (Nrp: $M = .71$, $SE = .02$; Rp−: $M = .64$, $SE = .02$), $t(59) = 2.78$, $p = .01$, $d = .40$. These findings replicate the standard effect of retrieval-induced forgetting. Contrary to the results of Jonker et al. (2013), however, no evidence of forgetting was observed in any of the restudy conditions. Specifically, in Experiment 1, Rp− items were not recalled differently than Nrp items in the no-context-change condition (Rp− items: $M = .70$, $SE = .02$; Nrp items: $M = .66$, $SE = .02$), $t(59) = 1.57$, $p = .12$, $d = .23$, the near-context-change condition (Rp− items: $M = .65$, $SE = .02$; Nrp items: $M = .66$, $SE = .02$), $t(59) = .31$, $p = .76$, $d = .05$, or the far-context-change condition (Rp− items: $M = .68$, $SE = .02$; Nrp items: $M = .66$, $SE = .02$), $t(59) = .91$, $p = .37$, $d = .11$. Again, the same pattern was observed in Experiment 2, such that Nrp items were not recalled differently than Rp− items in the no-context-change condition (Rp− items: $M = .69$, $SE = .03$; Nrp items: $M = .68$, $SE = .02$), $t(59) = .98$, $p = .33$, $d = .09$, the near-context-change condition (Rp− items: $M = .67$, $SE = .02$; Nrp items: $M = .67$, $SE = .02$), $t(59) = .28$, $p = .78$, $d = .00$, or the far-context-change condition (Rp− items: $M = .68$, $SE = .02$; Nrp items: $M = .70$, $SE = .02$), $t(59) = .91$, $p = .40$, $d = .10$.

Two 2 (Nrp vs. Rp−) × 3 (No-Context-Change vs. Near-Context-Change vs. Far-Context-Change) ANOVAs examining the three restudy conditions failed to reveal a significant interaction in Experiment 1, $F(1, 178) = 1.03$, $MSE = .02$,

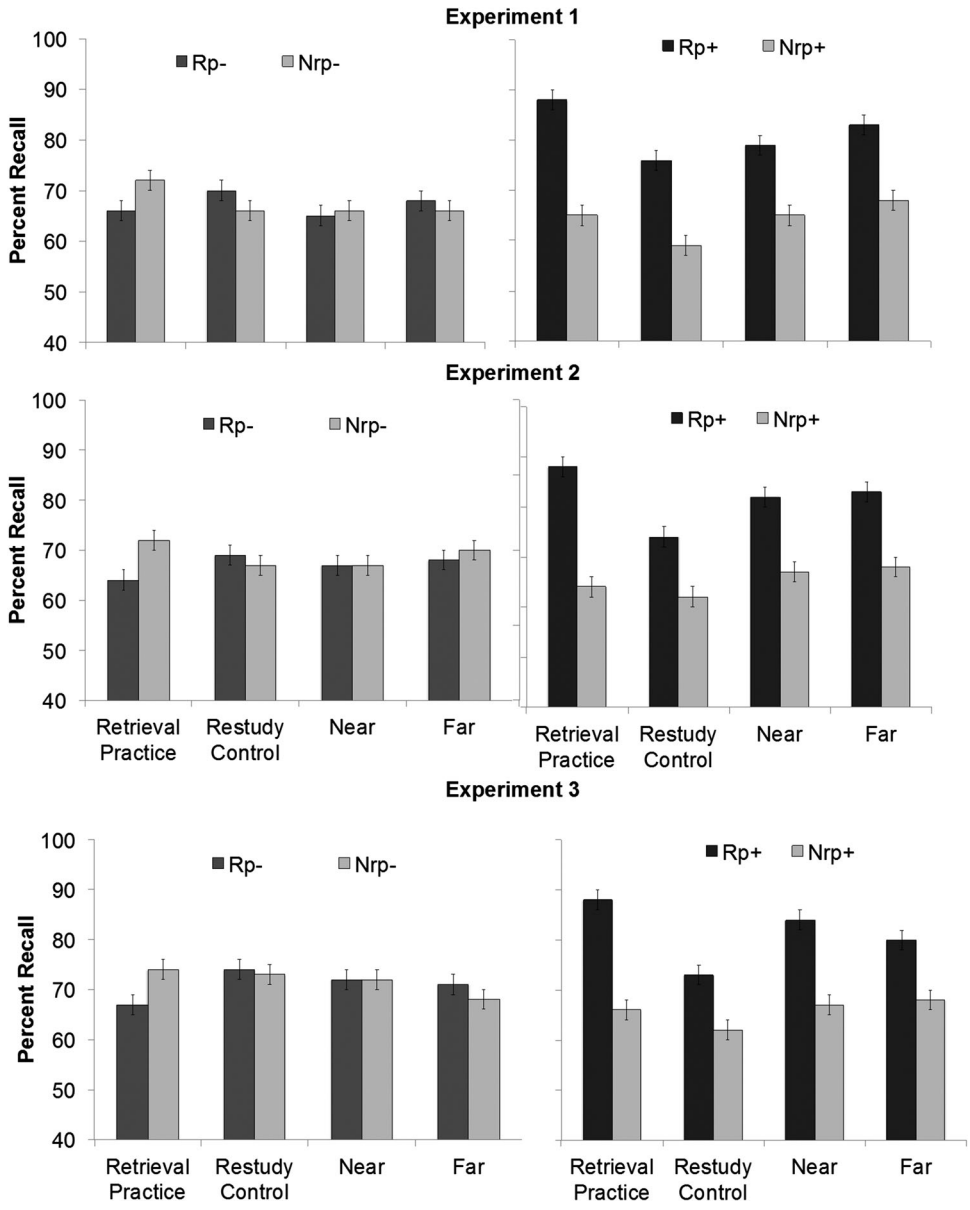

Figure 1. *Recall performance on the final test as a function of retrieval-practice status and experimental condition.*

$p = .36$, or Experiment, 2 $F(1, 178) = .75$, $MSE = .02$, $p = .47$, thus confirming that rates of forgetting did not differ significantly between the three restudy conditions.

One advantage to our design, compared to that of Jonker et al. (2013), is that we measured the consequences of retrieval practice and restudy within the same experiment, thereby allowing us to compare rates of forgetting in the two conditions directly. First, we examined whether retrieval practice elicited a greater forgetting effect than restudy without context change. A 2 (Nrp vs.

Rp–) × 2 (Retrieval Practice vs. No-Context-Change Restudy) ANOVA revealed a significant interaction such that the forgetting effect was significantly larger in the retrieval-practice condition than in the no-context-change restudy condition [Experiment 1: $F(1, 118) = 6.39$, $MSE = .02$, $p = .01$; Experiment 2: $F(1, 118) = 7.24$, $MSE = .02$, $p = .01$]. These results replicate the typical finding of retrieval-specificity and strength-independence that retrieval practice causes more forgetting than restudy practice (see e.g., Murayama et al., 2014).

Next, we examined whether retrieval practice caused significantly more forgetting than restudy even when restudy was accompanied by a context-change manipulation between study and restudy. For this analysis, we combined the near- and far-context-change conditions in each experiment to create overall context-change conditions. Two 2 (Nrp vs. Rp–) × 2 (Retrieval Practice vs. Context-Change Restudy) ANOVAs revealed significant interactions in both experiments, such that the forgetting effect observed in the retrieval-practice condition was significantly larger than that observed in the context-change condition [Experiment 1: $F(1, 178) = 4.02$, $MSE = .02$, $p < .05$; Experiment 2: $F(1, 178) = 4.60$, $MSE = .02$, $p = .03$]. Thus, contrary to the predictions of the context account, restudy with context change did not lead to as much forgetting as retrieval practice. Note that the interaction was also significant when we compared the retrieval-practice condition directly with the far-context-change condition [Experiment 1: $F(1, 118) = 4.48$, $MSE = .02$, $p = .04$; Experiment 2: $F(1, 118) = 6.65$, $MSE = .02$, $p = .01$].

EXPERIMENT 3

Experiments 1 and 2 attempted to extend Jonker et al.'s (2013) finding that an effect like retrieval-induced forgetting can be observed following restudy practice as long as such practice is accompanied by a mental-context change. Across both experiments, significant forgetting was observed in the standard retrieval-practice condition, but not in any of the restudy conditions. These findings are inconsistent with the predictions of the context account and suggest that a mental-context change between study and restudy practice is not sufficient to cause forgetting.

There are a number of small differences between the procedure we employed and that employed by Jonker and colleagues that could potentially account for the discrepancy in the results. For instance, Jonker et al. (2013) allotted 5 s for study and 10 s for restudy or retrieval practice, while we allotted 4 s and 7 s, respectively. Jonker and colleagues asked participants to study 6 categories comprised of 8 items each, while we used 8 categories comprised of 6 items each. Furthermore, while Jonker et al. (2013) asked participants to include a sketch of the layout of the home they were describing during the context change manipulation, we asked participants to only write descriptions of the home. Previous research has suggested that producing sketches can affect processes of retrieval and context establishment (e.g., Dando Wilcock, Milne, & Henry, 2009), so it seems possible that omitting this component contributed to the null effect of context change. To explore the potential role of some of these methodological differences, a third experiment was conducted in which we attempted to track more closely the procedures carried out by Jonker and colleagues.

Method

Participants and design

A total of 296 UCLA students participated for credit in an introductory psychology course. Retrieval-Practice Status (Nrp vs. Rp− vs. Rp+) was manipulated within subjects. Experimental condition (no-context-change restudy vs. no-context-change retrieval practice vs. near-context-change restudy vs. far-context-change restudy) was manipulated between subjects.

Materials

Six categories were selected, each consisting of eight high-frequency exemplars, for a total of 48 category–exemplar pairs (taken directly from Anderson et al., 1994). The pairs were

counterbalanced such that each item served equally often as an Rp+ item, Rp− item, and Nrp item.

Procedure

Participants were randomly assigned to one of the four conditions described below. All participants completed the experiment in a single room, accompanied by an experimenter to ensure that they complied appropriately with all instructions.

No-context-change restudy. This condition was identical to that implemented in Experiments 1 and 2, except that during the initial study phase, 48 exemplar pairs were presented via computer at a rate of one pair every 5 s. Once again, order was set randomly, with the constraint that no two consecutive pairs could be shown from the same category. Immediately following the completion of the study phase, participants were prompted to restudy half of the exemplars from half of the categories. Again, there were three rounds of practice such that the 12 restudied pairs were presented three times, but they were now presented for 10 s, resulting in a total of 36 restudy trials. After the restudy phase, and a 5-min distractor phase, a category-plus-one-letter-stem cued recall test was administered in which participants were given 6 s to recall each exemplar. Note that the testing procedure we implemented was block randomized, such that half of the Nrp items were tested along with the Rp− items in the first block, and the other half of the Nrp items were tested along with the Rp+ items in the second block. In contrast, Jonker and colleagues chose to block tested items by category.

Near-context-change restudy. This condition was the same as it was for Experiment 2, except that participants were asked to sketch the layout of their home as they mentally walked through each of the rooms. Participants were given one minute to complete this task.

Far-context-change restudy. This condition was identical to that of Experiment 2, except that participants were asked to sketch the layout of their parents' home as they mentally walked through each of the rooms. Participants were given one minute to complete this task.

No-context-change retrieval practice. This condition was identical to that implemented in Experiments 1 and 2, except that as participants attempted to retrieve each of the 12 to-be-practised exemplars 3 times, each category-plus-two-letter stem cue was presented for 10 s.

Results

Retrieval-practice performance

Participants recalled the appropriate exemplar on 89% of the retrieval-practice trials.

Final recall performance for practised items and baseline controls

As can be seen in Figure 1, a significant facilitation effect was observed in all four conditions, such that Rp+ items were recalled significantly better than were Nrp+ items (all $p < .001$).

Final recall performance for non-practised items from practised categories and baseline controls

Recall performance in for Rp− and Nrp− items on the final test is shown as a function of condition in Figure 1 and was analysed using a 2 (Rp− vs. Nrp–) × 4 (Near vs. Far vs. Retrieval Practice vs. Restudy) mixed ANOVA. Although the main effects of retrieval-practice status, $F(1, 292) = .40$, $MSE = .02$, $p = .53$, and context-change condition, $F(1, 292) = .14$, $MSE = .06$, $p = .93$, were not significant, a significant interaction was observed, $F(3, 292) = 4.52$, $MSE = .02$, $p < .05$.

As can be seen in Figure 1, forgetting was observed in the retrieval-practice condition, with Rp− items ($M = .67$, $SE = .02$) recalled significantly less well than Nrp items ($M = .73$, $SE = .02$), $t(73) = 2.97$, $p < .05$, $d = .34$. Once again, this finding replicates the standard effect of retrieval-induced forgetting. As was observed in Experiments 1 and 2, no evidence of forgetting was found in any of the restudy conditions. Specifically, Rp− items were not recalled differently than Nrp items in the no-context-change condition (Rp− items: $M = .73$, $SE = .02$; Nrp

items: $M = .71$, $SE = .02$), $t(73) = 1.31$, $p = .19$, $d = .12$, the near-context-change condition (Rp− items: $M = .72$, $SE = .02$; Nrp items: $M = .69$, $SE = .02$), $t(73) = 1.22$, $p = .23$, $d = .15$, or the far-context-change condition (Rp− items: $M = .73$, $SE = .02$; Nrp items: $M = .69$, $SE = .02$), $t(73) = 1.59$, $p = .12$, $d = .19$.

A 2 (Nrp vs. Rp–) × 3 (No-Context-Change vs. Near-Context-Change vs. Far-Context-Change) ANOVA examining the three restudy conditions failed to reveal a significant interaction, $F(1, 220) = .00$, $MSE = .02$, $p = .99$, thus confirming that rates of forgetting did not differ significantly between the three restudy conditions.

Once again, we examined whether retrieval practice elicited a greater forgetting effect than restudy without context change. A 2 (Nrp vs. Rp–) × 2 (Retrieval Practice vs. No-Context-Change Restudy) ANOVA revealed a significant interaction such that the forgetting effect was significantly larger in the retrieval-practice condition than in the no-context-change restudy condition, $F(1, 146) = 8.84$, $MSE = .02$, $p = .01$. These results replicate the typical finding of retrieval-specificity and strength-independence that retrieval practice causes more forgetting than restudy practice.

Next, we examined whether retrieval practice caused significantly more forgetting than restudy even when restudy was accompanied by a context-change manipulation between study and restudy. For this analysis, we combined the near- and far-context-change conditions in each experiment to create overall context-change conditions. The 2 (Nrp vs. Rp-) × 2 (Retrieval Practice vs. Context-Change Restudy) ANOVA revealed a significant interaction, such that the forgetting effect observed in the retrieval-practice condition was significantly larger than that observed in the context-change condition $F(1, 220 = 12.55$, $MSE = .02$, $p < .01$. Thus, once again, contrary to the predictions of the context account, restudy with context change did not lead to as much forgetting as retrieval practice. Note that the interaction was also significant when we compared the retrieval-practice condition directly with the far-context-change condition, $F(1, 146) = 10.50$, $MSE = .02$, $p < .01$.

GENERAL DISCUSSION

The context account clearly predicts that forgetting should have emerged in the context-change restudy conditions, and possibly to a greater extent in the far condition than in the near condition. Across three experiments and a total of 776 subjects, however, significant forgetting was only observed in the retrieval-practice condition. These results are difficult to reconcile with the core tenets of the context account of retrieval-induced forgetting. If it was the change in context induced by retrieval practice that was responsible for Rp− items becoming less recallable than Nrp items in the retrieval-practice condition, then such an effect should also have been observed in the restudy condition when a change in context was induced via experimental manipulation.

These results are consistent, however, with predictions of the inhibitory account of retrieval-induced forgetting (i.e., strength independence and retrieval specificity). That is, one line of evidence that has provided compelling support for the inhibitory account is that retrieval-induced forgetting is often observed only when information is actively retrieved from memory in the face of competition from related contextually inappropriate information (Anderson, 2003; Murayama et al., 2014; Storm & Levy, 2012). Restudy practice would not be expected to produce forgetting, even if preceded by a contextual shift, because there would have been no need to inhibit the non-practised items from the restudied categories. The present results thus reaffirm evidence of strength independence and retrieval specificity and suggest that such effects cannot be explained by the context-account of retrieval-induced forgetting.

It is somewhat unclear why our results differ so strikingly from those of Jonker and colleagues—that is, why did Jonker et al. observe a significant forgetting effect following restudy accompanied by mental-context change, whereas we did not? One possibility is that their finding reflected a false positive. Indeed, Jonker et al. observed a small to medium effect size in both the initial experiment and the replication ($d = .40$ and $d = .51$) with a combined sample of only 60 participants across

experiments. In comparison, collapsing across all of the restudy-plus-context change conditions of the three experiments reported here, our sample consisted of 388 participants. Even when we combined across all of these participants, however, we still failed to observe any evidence of forgetting (Rp− items: $M = .69$, $SE = .01$; Nrp items: $M = .68$, $SE = .01$), $t(387) = -1.165$, $p = .25$, $d = .05$ A power analysis based on the effect size observed by Jonker and colleagues suggests that a sample of this size should have been more than sufficient to observe an effect (power = .99).

To examine the data further, Bayes factors were estimated using Bayesian Information Criteria (Wagenmakers, 2007). This model compares the fit of the data under the null hypothesis and the alternative hypothesis and computes the probability that the alternative is better than the null, or vice versa. All of the analyses discussed below are one-tailed and assume that evidence for the alternative hypothesis will come in the form of a positive effect of retrieval-induced forgetting (Nrp > Rp–). The size of the effect is presumed to reflect about a 5.0% mean difference between Nrp and Rp–, as was observed in a recent meta-analysis of retrieval-induced forgetting for studies that effectively controlled for output interference (Murayama et al., 2014).

When the data were collapsed across all three experiments, in the standard retrieval-practice condition an estimated Bayes factor suggested that the data were 6685.63 times more likely to occur under a model including an effect of retrieval-induced forgetting than under a model without it. Hence, the data were strongly in favour of the alternative hypothesis. These results are to be expected, according to both the inhibitory account and the contextual-cuing account. Similarly, collapsed across all three experiments, in the restudy, no context change condition an estimated Bayes Factor suggested that the data were 11.37 times more likely to occur under the null hypothesis, or within a model that does not include an effect of retrieval-induced forgetting. Once again, these results are to be expected according to both the inhibitory account and the contextual-cuing account. Critically, when the data were collapsed across both restudy-plus-context-change-conditions, and across all three experiments, an estimated Bayes Factor suggested that the data were 11.28 times more likely under the null hypothesis, or within a model that does not include an effect of retrieval-induced forgetting. This index provides strong evidence in favour of the null hypothesis and can be directly compared to the results obtained by Jonker et al. (2013). When the results of Experiment 2a were combined with the replication reported in Jonker et al., for a total of 60 participants, an Estimated Bayes Factor of 155.83 was obtained. This suggests that, unlike our own data, the data obtained by Jonker et al. were far more likely to occur under a model that includes an effect of retrieval-induced forgetting. When we combined the results of our three experiments with the results of the two experiments conducted by Jonker et al., however, the Bayes Factor suggested that the data were 5.14 times more likely to occur under the null hypothesis than under a model that includes an effect of retrieval-induced forgetting. Thus, when all of the data currently available are combined and analysed together, there is still evidence favouring the null hypothesis.

These discrepant results are somewhat puzzling, and it remains unclear why an effect of context-change-induced forgetting would emerge in one case and not the other. It should be noted that although our procedure was highly similar to that adopted by Jonker et al., there were some subtle methodological differences. For example, during the restudy phase participants in Jonker et al.'s experiments read each category–exemplar pair aloud for 5 s, while participants in each of the experiments reported above were instructed to silently restudy each pair for 10 s. Also, while participants in Jonker et al.'s studies were required to engage in a sematic generation task (i.e. generating a list of countries) as a distractor following the restudy phase, our participants were asked to complete a series of complex multiplication problems. It is unclear whether these minor differences in procedure contributed to the discrepant pattern of results. In any case, transparency regarding these differences should inform future research investigating any critical factors that must be present in order for effects of context to reliably emerge.

One other notable methodological difference between the present studies and those reported by Jonker et al. is the nature of the final test. Recall that the final test for each experiment reported here was divided into two test blocks: Rp− items and half of the Nrp items (Nrp− items) were tested in the first block, and Rp+ items and the other half of the Nrp items (Nrp+ items) tested in the second block. In contrast, the final test in Jonker et al. (2013, Experiment 2a, and its subsequent replication) was blocked by category, such that participants first outputted the Rp− items of a given category, followed by the Rp+ items of that same category. This procedure was followed until all of the categories had been tested. One important feature of this blocked-by-category testing procedure as compared to that employed in the present study is that participants may have been less likely to target the study context when attempting to recall Rp− items. Perhaps, for example, being tested on Rp+ items leads participants to become more likely to target the inappropriate retrieval practice context when attempting to recall Rp− items from other categories. In contrast, because all of the Rp− items were tested prior to Rp+ items in the current study, subjects may have been more likely to successfully reinstate (and maintain) the study context when attempting to recall Rp− items. This is all just speculative of course, and future research is needed to address this possibility.

Although the present results may seem surprising in light of Jonker and colleagues' results, much of the data in the extant literature is at odds with the predictions made by the contextual-cuing account. For example, when a forget instruction is interpolated between study and retrieval-practice, retrieval-induced forgetting is not observed (Storm, Bjork, & Bjork, 2007). Forget instructions have been shown to reliably induce a change in mental context (Sahakyan & Kelley, 2002), and thus the contextual-cuing account should predict substantial retrieval-induced forgetting; the results however, suggest otherwise. Similarly, retrieval-induced forgetting has been found to be reduced or eliminated when a negative mood or stress is induced prior to retrieval practice (e.g., Bäuml & Kuhbandner, 2007; Koessler, Engler, Riether, & Kissler, 2009). Both of these manipulations should have led to significant changes in internal context, yet no retrieval-induced forgetting was observed. Moreover, retrieval-induced forgetting has also been shown to be eliminated when participants perform retrieval practice under divided attention (Román, Soriano, Gómez-Ariza, & Bajo, 2009). The contextual-cuing account contends that retrieval practice causes forgetting, whereas restudy does not, because the more difficult retrieval task leads to a shift in context. Presumably, combining the retrieval-practice task with a concurrent updating task should have, if anything, enhanced the shift in context, yet once again no forgetting was observed.

While all of these results contradict many of the fundamental assumptions of the context account, they are entirely consistent with predictions made by the inhibitory account. For instance, interpolating a forget instruction between study and retrieval practice is presumed to reduce retrieval-induced forgetting because it reduces the extent to which non-practised items cause associative interference during retrieval practice. Similarly, retrieval-induced forgetting is presumed to be diminished when concurrent tasks are performed during retrieval practice because both tasks require inhibitory control, and thus fewer resources are available to be allocated toward suppressing competing items during retrieval practice.

Taken together, the results of the current study, combined with the results of studies like those reviewed above, suggest that inserting a mental context change between study and restudy is not sufficient to produce retrieval-induced forgetting. This is not to say that the context account, or some version of it, is entirely incorrect but, rather, that one of the critical tenets of the account is not well supported. There is still some evidence supporting the context account that we did not address in the current research. While implementing the standard retrieval-practice paradigm, for example, Jonker and colleagues found that reinstating the study context at test eliminated retrieval-induced forgetting, whereas reinstating the retrieval-practice context did not. This finding suggests that contextual cues at test can influence the extent to which selective practice causes forgetting—a possibility

that certainly warrants further theoretical consideration and empirical investigation.

REFERENCES

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, *49*(4), 415–445. doi:http://dx.doi.org/10.1016/j.jml.2003.08.006

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1063–1087. doi:http://dx.doi.org/10.1037/0278-7393.20.5.1063

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, *7*(3), 522–530. doi:http://dx.doi.org/10.3758/BF03214366

Bäuml, K. H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*, *13*(4), 356–360.

Bäuml, K.-H., & Kuhbandner, C. (2007). Remembering can cause forgetting — but not in negative moods. *Psychological Science*, *18*(2), 111–115. doi:10.1111/j.1467-9280.2007.01857.x

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 123–144). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309–330). Hillsdale, NJ: Erlbaum.

Dando, C., Wilcock, R., Milne, R., & Henry, L. (2009). A modified cognitive interview procedure for frontline police investigators. *Applied Cognitive Psychology*, *23*(5), 698–716.

Delaney, P. F., Sahakyan, L., Kelley, C. M., & Zimmerman, C. (2010). Remembering to forget: The amnesic effect of daydreams. *Psychological Science*, *21*(7), 1036–1042. doi:10.1177/0956797610374739

Jonker, T. R., Seli, P., & MacLeod, C. M. (2013). Putting retrieval-induced forgetting in context: An inhibition-free, context-based account. *Psychological Review*, *120*(4), 852–872.

Koessler, S., Engler, H., Riether, C., & Kissler, J. (2009). No retrieval-induced forgetting under stress. *Psychological Science*, *20*(11), 1356–1363. doi:http://dx.doi.org/10.1111/j.1467-9280.2009.02450.x

Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: A meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin*, *140*, 1383–1409.

Raaijmakers, J. G. W., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the inhibition theory for forgetting. *Journal of Memory and Language*, *68*(2), 98–122. doi:http://dx.doi.org/10.1016/j.jml.2012.10.002

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*(3), 181–210. doi:10.1111/j.1745-6916.2006.00012.x

Román, P., Soriano, M. F., Gómez-Ariza, C. J., & Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychological Science*, *20*(9), 1053–1058. doi:http://dx.doi.org/10.1111/j.1467-9280.2009.02415.x

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064–1072.

Smith, S. M., & Manzano, I. (2010). Video context-dependent recall. *Behavior Research Methods*, *42*, 292–301.

Storm, B. C., Angello, G., Buchli, D. R., Koppel, R. H., Little, J. L., & Nestojko, J. F. (2015). A review of retrieval-induced forgetting in the contexts of learning, eye-witness memory, social cognition, autobiographical memory, and creative cognition. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 141–194). San Diego, CA: Academic Press/Elsevier Inc.

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2007). When intended remembering leads to unintended forgetting. *The Quarterly Journal of Experimental Psychology*, *60*(7), 909–915. doi:http://dx.doi.org/10.1080/17470210701288706

Storm, B. C., & Levy, B. J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition*, *40*(6), 827–843. doi:10.3758/s13421-012-0211-7

Verde, M. F. (2012). Retrieval-induced forgetting and inhibition: A critical review. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 47–80). San Diego, CA: Academic Press/Elsevier Inc.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*, 779–804.