4-2013

Providing Corrective Feedback During Retrieval Practice Does Not Increase Retrieval-Induced Forgetting

Matthew R. Erdman
Iowa State University

Jason C.K. Chan
Iowa State University, ckchan@iastate.edu

Providing Corrective Feedback During Retrieval Practice Does Not Increase Retrieval-Induced Forgetting

Abstract

Recalling a subset of studied materials can impair subsequent retrieval of related, nontested materials. In two experiments, we examined the influence of providing corrective feedback (no feedback, immediate feedback, delayed feedback) during retrieval practice on this retrieval-induced forgetting effect. Performance was assessed with category cued recall (e.g., recall all exemplars studied under Weather), category-and-stem cued recall (e.g., Weather–B____), and recognition. We report a dissociation between the effects of feedback on memory of the tested materials and the nontested materials. Whereas providing immediate or delayed feedback (compared to no feedback) improved recall and recognition of the tested items, it had no influence on retrieval-induced forgetting. These results are consistent with the inhibition account of retrieval-induced forgetting. From an applied perspective, this finding is encouraging for students and educators who use testing to foster learning.

Keywords

Disciplines

Cognitive Psychology | Educational Psychology

Comments

Providing Corrective Feedback During Retrieval Practice Does Not Increase

Retrieval-Induced Forgetting

Matthew R. Erdman

Jason C.K. Chan

Iowa State University

(Word Count including References: 6,217)

Please address correspondence to:

Jason C.K. Chan
Department of Psychology
Iowa State University
Ames, IA 50011
USA
515-294-6946
ckchan@iastate.edu

Abstract

Recalling a subset of studied materials can impair subsequent retrieval of related, nontested materials. In two experiments, we examined the influence of providing corrective feedback (no feedback, immediate feedback, delayed feedback) during retrieval practice on this retrieval-induced forgetting effect. Performance was assessed with category cued recall (e.g., recall all exemplars studied under Weather), category-and-stem cued recall (e.g., Weather – B___), and recognition. We report a dissociation between the effects of feedback on memory of the tested materials and the nontested materials. Whereas providing immediate or delayed feedback (compared to no feedback) improved recall and recognition of the tested items, it had no influence on retrieval-induced forgetting. These results are consistent with the inhibition account of retrieval-induced forgetting. From an applied perspective, this finding is encouraging for students and educators who use testing to foster learning.

(Word Count: 135)

Providing Corrective Feedback During Retrieval Practice Does Not Increase

Retrieval-Induced Forgetting

In recent years, researchers have identified retrieval practice, or testing, as a powerful

tool to improve learning (Roediger & Karpicke, 2006), and providing feedback (i.e., the correct

answer) can further augment the benefits of retrieval practice (Kang, McDermott, & Roediger,

2007). Although much research has demonstrated that the combination of testing and corrective

feedback can dramatically enhance retention of the tested material (e.g., McDaniel, Howard, &

Einstein, 2009), its impact on retention of the ***nontested material*** remains unknown. In the

present investigation, we began to answer this question.

Substantial work has investigated the effects of retrieval practice on subsequent memory

of the nontested material. In general, performing retrieval practice on a subset of the studied

items can impair later recall of the items that did not receive retrieval practice – a finding termed

retrieval-induced forgetting (Anderson, 2003; Butler, 2010; but see also Chan, 2009). In a

typical study of its kind, participants study category-exemplar word pairs (e.g., *Fruit* – Orange,

Spice – Nutmeg, *Spice* – Thyme, *Weather* – Drought) and then perform retrieval practice on

some of the exemplars from half of the studied categories. This retrieval practice phase produces

three types of items: ***tested*** (i.e., tested items from the practiced categories), ***nontested*** (i.e.,

nontested items from the practiced categories), and ***control*** (i.e., items from categories in which

no exemplars were practiced).[1] Following a delay, participants' memory for all of the exemplars

---

[1] In the retrieval-induced forgetting literature, the tested, nontested, and control items are often

referred to as Rp+, Rp-, and Nrp, respectively. We opted to use non-abbreviated terms here

because we feel that their meanings are more transparent.

is tested. Typically, the nontested items exhibit poorer memory than the control items. This

finding has generated considerable research interest (for a recent review, see Storm & Levy,

2012). Curiously, despite the extensive literature on both retrieval-induced forgetting and

corrective feedback, no study has examined the influence of feedback during retrieval practice on

subsequent recall of the nontested materials. In fact, to the best of our knowledge, only two

studies in the vast literature on retrieval-induced forgetting have implemented corrective

feedback during retrieval practice (Ciranni & Shimamura, 1999; Raaijmakers & Jakab, 2012;

Roediger & Karpicke, 2006), but neither manipulated the administration of feedback as a

variable (i.e., a control, no-feedback condition was not included).

Researchers have advocated implementation of testing (Kang et al., 2007; Roediger &

Karpicke, 2006) and feedback (Butler, Karpicke, & Roediger, 2008; Kang et al., 2007; e.g.,

McDaniel et al., 2009) to enhance learning in the classroom. Because it is unrealistic to expect

students to test themselves on all of the study material, it is important to understand the effects of

retrieval practice and feedback on memory for the nontested material. Equally important is that

data from the present study have implications for the theoretical understanding of retrieval-

induced forgetting.

Broadly speaking, two theories have been proposed to explain retrieval-induced

forgetting. The interference theory suggests that retrieval-induced forgetting occurs because

performing retrieval practice on some of the studied items selectively strengthens these items in

memory. When one needs to recall the nontested items later, the stronger, more accessible tested

items block retrieval of the weaker, nontested items (Anderson, 2003; Butler, 2010; Camp,

Pecher, Schmidt, & Zeelenberg, 2009; but see also Chan, 2009; Raaijmakers & Jakab, 2013;

Williams & Zacks, 2001). That is, selective retrieval practice of a subset of the studied items

creates variations in item strength within a study list, with some items (the tested ones) being

much more accessible than others (the nontested ones). In this sense, retrieval-induced

forgetting is akin to the list-strength effect, where increasing the strength of some items in a list

impairs recall of the remaining items (Tulving & Hastie, 1972). Based on this idea, providing

feedback should ***exacerbate retrieval-induced forgetting*** because feedback would further

strengthen the tested exemplars and their association with the category cue, thus exacerbating the

likelihood that the tested items would interfere with retrieval of the nontested items.

   Another explanation for retrieval-induced forgetting is based on the idea that competition

during retrieval practice elicits inhibition of the nontested materials (Anderson, 2003; Storm &

Levy, 2012). Unlike interference theory, inhibition theory would lead one to predict no

influence of feedback on retrieval-induced forgetting. Because providing feedback ***after*** the

retrieval attempt would not alter the level of retrieval competition (and therefore inhibition), it

should not affect the magnitude of retrieval-induced forgetting. Moreover, one prominent

assumption of inhibition theory is that retrieval suppression is independent of the strength or

retrieval success of the tested items (Anderson, 2003; Bauml, 2002; Storm & Nestojko, 2010;

Storm, Bjork, Bjork, & Nestojko, 2006). The reasoning is that when one attempts to recall a

target memory, other items in memory (e.g., the items that share the retrieval cue with the target)

compete for retrieval. To resolve this competition, one must suppress the representation of these

nontarget items. The critical assumption of inhibition theory is that varying the strength of the

target items (and by extension, varying whether it is possible to retrieve the target, Storm &

Nestojko) should not alter the level of retrieval-induced forgetting observed, because inhibition

is triggered by competition initiated by *the nontarget items* (Anderson, Bjork, & Bjork, 2000; see

the competition-dependence assumption of inhibition, 1994). Therefore, according to inhibition theory, providing feedback should have no impact on retrieval-induced forgetting.

In the present study, we tested the influence of feedback on retrieval-induced forgetting by varying the presence and timing of feedback administration (immediate vs. delayed). Previous research has shown that delaying feedback can sometimes increase its benefits (Butler, Karpicke, & Roediger, 2007). If delayed feedback further strengthens memory for the tested items relative to immediate feedback, then it should further increase the magnitude of retrieval-induced forgetting according to the interference account. Specifically, some participants received feedback immediately after each retrieval attempt (i.e., ***immediate feedback***), whereas others received feedback only at the conclusion of the retrieval practice phase (i.e., ***delayed feedback***).

In Experiment 1, we attempted to maximize the potential contribution of interference to retrieval-induced forgetting by employing a category cued recall test (if feedback does increase the level of retrieval-induced forgetting, it would likely be detected here). In Experiment 2, we reduced the contribution of interference by administering a recognition test (we describe the motivation and reasoning behind this experiment more thoroughly following presentation of the results from Experiment 1). In sum, we sought to examine the effects of providing feedback during retrieval practice on subsequent memory of the nontested material. The experiments used a 3 (feedback condition: no feedback, immediate feedback, delayed feedback) X 3 (item type: tested, nontested, control) mixed design. Feedback, operationally defined here as providing the correct answer, was manipulated between-subjects and item type was manipulated within-subjects.

Experiment 1

Method

      Participants. Ninety-six undergraduate students at Iowa State University participated for research credit, with 32 in each feedback condition.

      Materials. Participants studied 48 category-exemplar word pairs. There were eight categories and six exemplars in each category. The word pairs were generated using the following rules: all category names and exemplars were single words, the category typicality rating for all exemplars were below .40 (range = .01-.38, M = .12, Battig & Montague, 1969), and the exemplar words had a low to medium word frequency (*range* = 1-101, *M* = 34.81, Thorndike & Lorge, 1942). See the Appendix for the materials.

      Procedure. Participants were tested on individual computers in groups of up to eight. During the encoding phase, each word pair appeared for 5 s with a 500 ms interstimulus interval. The pairs were presented in a random order, but none in the same category were presented consecutively. Following the study phase, participants performed retrieval practice on half (three) of the exemplars from half (four) of the categories, producing the three item types: tested, nontested, and control. The pairs that received retrieval practice were counterbalanced across participants. During each retrieval practice trial, participants were given the category name and a two-letter stem (e.g., *Weather* – Bl_____) and were instructed to type in the entire exemplar (i.e., Blizzard). Participants performed the retrieval practice phase three times, with a fresh random order for each, but no exemplars from the same category were tested on consecutive trials. In the ***no feedback*** condition, participants had 7 s to recall the target in each retrieval practice trial. In the ***immediate feedback*** condition, participants had 6 s to recall the target, and the correct answer (the entire cue-target pair) appeared for 1 s immediately afterward. In the ***delayed feedback*** condition, participants were given 6 s to recall each target. At the end of the

entire retrieval practice phase, the correct answers (i.e., the cue-target pairs) were shown for 3 s

each and in a random order. Participants were not aware that they would receive feedback until

the end of the retrieval practice phase, and the feedback was shown once. We decided not to

inform participants about the impending feedback here because the delayed feedback condition

was designed to strengthen accessibility of the tested items without possibly altering retrieval

effort during the retrieval practice phase (i.e., knowing that one would receive corrective

feedback may make one devote less effort to the retrieval practice task). Following the retrieval

practice phase, all participants performed the operation span task (Unsworth, Heitz, Schrock, &

Engle, 2005), followed by the videogame Tetris, for a total of 20 min.

During the final test phase, participants first completed a ***category cued recall test***. For

each trial, a category name appeared on the top of the screen, and participants had 30 s to type in

all of its studied exemplars. The order of the categories was randomized for each participant.

Afterwards, we administered a ***category-and-stem cued recall test*** (e.g., "*Spice* – N____").

Participants had 7 s to complete each trial. The categories were presented in a random order.

Exemplars within each category were presented consecutively and in a random order, but the

nontested items always appeared before the tested items (to reduce output interference, Anderson,

2003). We elected not to report results from this test because they were contaminated by the

category cued recall test. For the purpose of disclosure, however, we note that the data from this

test were completely consistent with those from the category cued recall test.

Results

Statistical significance was determined by an alpha level of .05 (two-tailed). Partial eta

squared (η_p^2) indicates effect size for analysis of variance (ANOVA), and Cohen's *d* indicates

effect size for *t*-tests.

Results from the Retrieval Practice Phase. A 3 (retrieval practice trial: 1, 2, 3) X 3

(feedback condition: none, immediate, delayed) ANOVA showed that recall probability rose

over the three retrieval practice trials, $F(2, 186) = 72.67$, $\eta_p^2 = .61$. However, this main effect

was moderated by an interaction between retrieval practice trial and feedback condition, $F(4,$

$186) = 19.02$, $\eta_p^2 = .29$. Whereas performance showed modest gains across the retrieval practice

trials in the no feedback (Ms = .70, .75, .75 in the first, second and third retrieval practice trials,

respectively) and delayed feedback conditions (Ms = .67, .71, .75), its improvement was, as

expected, far greater in the immediate feedback condition (Ms = .61, .90, .97).

Results from the Final Test Phase. Results from the category cued recall test are

displayed in Figure 1. A 3 (feedback condition) X 3 (item type) ANOVA revealed a main effect

of item type, $F(2, 92) = 376.01$, $\eta_p^2 = .89$. This main effect was characterized by a significant

retrieval-induced forgetting effect (i.e., poorer performance for the nontested items relative to the

control items), $t(95) = 8.46$, $d = .97$, and a testing effect (i.e., better performance for the tested

items relative to the control items), $t(95) = 19.27$, $d = 2.10$. Feedback condition did not produce

a main effect, $F = 1.87$, $p = .16$, but its marginal interaction with item type showed that feedback

affected recall of only the tested items, $F_{interaction}(4, 186) = 2.33$, $p = .06$. Specifically, both

delayed feedback ($M = .75$) and immediate feedback ($M = .74$) promoted recall of the tested

items over no feedback ($M = .62$), ts > 2.15, ds > .53. Most important for present purposes,

however, is that feedback did not increase the magnitude of retrieval-induced forgetting (.15 with

no feedback, .12 with immediate feedback, .15 with delayed feedback). This finding is

consistent with the prediction based on the inhibition account.

One potential concern with our data is that recall probability of the nontested items was

quite low in the no feedback condition (.16), which might have masked any detrimental effects

of providing feedback on recall of the nontested items. To address this possibility, we reanalyzed our data based on the 16 participants (i.e., the top half) with the highest nontested items recall probability in each feedback condition (overall $M = .29$). Even in this analysis, providing feedback did not exacerbate retrieval-induced forgetting (.09 for no feedback, .10 for immediate feedback, .05 for delayed feedback).

Another possible concern is that our feedback effect was modest. That is, perhaps our feedback manipulation did not strengthen the tested items enough to affect retrieval-induced forgetting. To address this possibility, we examined whether the magnitude of retrieval-induced forgetting increased for participants who benefited the most from corrective feedback. It is possible to estimate the effectiveness of feedback for participants in the immediate feedback condition on an individual basis by comparing recall performance of the third retrieval practice trial with that of the first. Because recall performance was relatively stable across the retrieval practice trials without between-trial feedback (see the no feedback and delayed feedback conditions), it is reasonable to assume that participants who demonstrated the greatest gains in performance across the retrieval practice trials were the ones who benefitted the most from feedback. We performed a median split analysis based on the difference in performance between retrieval practice trial 1 and trial 3. Remarkably, nearly identical levels of retrieval-induced forgetting (.12 and .13, respectively) was observed between participants who benefitted the most (.49) from feedback and those who benefitted relatively little (.25). Thus, the size of the retrieval-induced forgetting effect appears to be independent of the size of the feedback benefit.

An output order analysis was conducted to examine whether feedback affected output interference, which could mask the effects of retrieval inhibition (Anderson, 2003). We divided the average output positions of the tested and nontested items by the total number of items

recalled per category for each participant. For example, if a participant recalled five items from

a category and the first three recalled items were tested, the average output position for the tested

items would be $[(1 + 2 + 3)/3]/5 = .40$, whereas the average output position for the nontested

items would be $[(4 + 5)/2]/5 = .90$. To facilitate comparison across feedback conditions, we

divided the output position of the nontested items by that of the tested items to yield a single

score of output ratio (e.g., $.90/.40 = 2.25$). If this ratio is greater than 1, then the nontested items

were output later than the tested items (which was expected). Of importance here is whether the

relative output positions of the tested and nontested items were influenced by feedback – they

were not, as the output position ratio was 1.36 for the no feedback condition, 1.53 for the

immediate feedback condition, and 1.45 for the delayed feedback condition, $F < 1, p = .43$.

Experiment 2

The results from Experiment 1 show that feedback does not influence the magnitude of

retrieval-induced forgetting. Although this finding is consistent with the inhibition account, a

lingering possibility is that different factors contributed to retrieval-induced forgetting depending

on whether or not participants received feedback. Specifically, providing feedback during

retrieval practice could have eliminated inhibition by reducing retrieval competition, but a

retrieval-induced forgetting effect was nonetheless observed because feedback differentially

strengthened the tested items, causing these items to block recall of the nontested items during

the category cued recall test.[2] To address this concern, we administered a yes/no recognition test

in Experiment 2. Because the studied category cue is not presented in free choice recognition,

---

[2] Note that this explanation has trouble accounting for the equivalent performance between the

immediate feedback and delayed feedback conditions.

this test should significantly weaken the possibility that the category cue would trigger retrieval

of the tested items, which in turn blocks retrieval of the nontested items (Aslan & Bäuml, 2011;

Veling & van Knippenberg, 2004). However, Verde (2012) argued that recognition is not

immune to interference effects because recollection (a recall-like retrieval process) can

contribute to its performance (see also Raaijmakers & Jakab, 2013). More concretely,

interference effects can operate in a recognition test if participants spontaneously recall the

category cue when the target is presented (Camp et al., 2009), which activates memory of the

tested items, which in turn block retrieval of the nontested target (despite it being presented as

the retrieval cue in a recognition test). Notably, even if retrieval-induced forgetting in a

recognition test does not represent unequivocal evidence for inhibition, two important factors

remain. First, it is well known that, compared to recall, recognition is far less susceptible to

interference effects in general (e.g., Ratcliff, Clark, & Shiffrin, 1990; Yonelinas, Hockley, &

Murdock, 1992). Second, even if recognition is susceptible to retrieval blocking, interference

theory would still predict that feedback should increase retrieval-induced forgetting. Because

feedback strengthens the association between the category cue and the tested items, it should

increase the likelihood that covert retrieval of the category cue would then activate the tested

items, thus enhancing any putative blocking effects. To be clear, if the retrieval-induced

forgetting effect in Experiment 1 was driven by inhibition, results from Experiment 2 should

mirror those from Experiment 1 – that is, feedback would have no influence on retrieval-induced

forgetting.

Method

        Participants. A total of 144 undergraduate students at Iowa State University participated

for research credit, with 48 participants in each feedback condition.

Materials and Procedure. Because the recognition final test required a set of foils, we

selected 48 additional exemplars. The category typicality criterion was expanded to acquire the

necessary items. The category "Religion" was replaced by "Metal" because we were unable to

obtain enough exemplars for the former category. Typicality of all exemplars was less than .50

(*range* = .01-.46, *M* = .11, *SD* = .09). Exemplars were also selected based on Thorndike and

Lorge's (1944) norms (*range* = 1-226, *M* = 53.38, *SD* = 52.38) and the Kucera and Francis

(1967) norms (*range* = 1-15, *M* = 3.75, *SD* = 3.38). See the Appendix for the materials.

Whether a word served as a target or a lure was counterbalanced across participants. The

procedure of Experiment 2 was identical to Experiment 1 except that participants completed a

yes/no recognition test. The test was self-paced and contained 96 trials (half studied and half

new) presented in a random order. Category names were never presented during the recognition

test.

Results

Results from the Retrieval Practice Phase. Similar to Experiment 1, a 3 (retrieval

practice trial) X 3 (feedback condition) ANOVA revealed that both feedback and retrieval

practice trials produced significant main effects, $Fs > 12.92$, $\eta_p^2 > .22$. More importantly,

feedback interacted with retrieval practice trials, $F(4, 282) = 18.77$, $\eta_p^2 = .21$. As expected,

recall probability rose substantially across the retrieval practice trials (*Ms* = .68, .90, .93) in the

immediate feedback condition but not in the delayed feedback (*Ms* = .69, .73, .74) and no

feedback conditions (*Ms* = .64, .66, .66).

Results from the Final Test Phase. False alarm rates did not differ significantly based

on feedback (no-feedback = .12, immediate feedback = .13, delayed feedback = .16), $F = 1.31$, p

= .27. We conducted all analyses based on *d'* (see Figure 2 for means). Whenever the hit rate

was 1 and the false alarm rate was 0, we applied the standard correction method whereby the

false alarm rate was converted to $1/2N$ (where N was the number of trials) and the hit rate was

converted to $1 – 1/(2N)$ (Macmillan & Creelman, 2005).

A 3 (feedback condition) X 3 (item type) ANOVA revealed a main effect of item type,

$F(2, 142) = 289.20$, $\eta_p^2 = .81$. There was a significant retrieval-induced forgetting effect, $t(143)$

$= 4.30$, Cohen's $d = .23$, and a testing effect, $t(143) = 20.61$, Cohen's $d = 1.33$. Feedback

condition produce a marginally significant main effect, $F(2, 141) = 2.56$, $p = .08$, and it

interacted with item type, $F(4, 282) = 3.99$, $p = .04$. This interaction was driven by the fact that

feedback affected recognition performance of only the tested items. Specifically, both delayed

feedback ($d' = 2.93$) and immediate feedback ($d' = 2.91$) enhanced recognition of the tested

items over no feedback ($d' = 2.40$), ts > 3.12, Cohen's $d > .64$. However, feedback had virtually

no influence on the magnitude of retrieval-induced forgetting in d' (.19 with no feedback, .19

with immediate feedback, .17 with delayed feedback).

Again, we performed a median split analysis to examine whether participants who

benefitted the most from feedback (as estimated by the difference of recall probability between

retrieval practice trial 1 and trial 3 = .39) would show greater retrieval-induced forgetting than

participants who benefitted the least (.13) – they did not. Rather, the two groups showed similar

levels of retrieval-induced forgetting in d' (.18 for high feedback benefit and .21 for low

feedback benefit), $t < 1$, $p = .82$. In sum, similar to Experiment 1, despite feedback increasing

performance of the tested items, it did not exacerbate retrieval-induced forgetting.

General Discussion

In two experiments, we attempted to address whether obtaining corrective feedback

during retrieval practice would exacerbate the magnitude of retrieval-induced forgetting.

Overall, feedback increased recall and recognition probabilities of the tested items without impairing those of the nontested items. From a theoretical perspective, these findings are consistent with the inhibition account. From an applied perspective, they suggest that acquiring feedback during retrieval practice enhances memory of the tested materials without impairing memory of the nontested materials beyond that of receiving retrieval practice alone.

Theoretical Implications

The finding that feedback increased performance of the tested items without impairing that of the nontested items is consistent with predictions based on the inhibition account. A possible concern with this conclusion is that support for inhibition rests on a null effect, and perhaps our experiments were simply not sensitive enough to detect the effects of feedback on retrieval-induced forgetting. We do not find this argument compelling for several reasons. Although feedback failed to increase the size of retrieval-induced forgetting, we were able to detect its facilitative effects on the tested items in two experiments using different retrieval tasks. Moreover, we presented the entire category-exemplar pair as feedback. This procedure ensured that the cue-target association, rather than just the target, was strengthened by feedback. This is important because strengthening the target representation alone (instead of strengthening the cue-target association) may not be sufficient to produce interference effects in recall (Verde, 2009).

In addition, the final test in Experiment 1 was category cued recall, which should provide abundant opportunities for interference effects to be revealed (Bauml, 1997). Unlike tests in which output order can be manipulated, a category cued recall test allows participants to output items in any order. Because the tested items are likely to be recalled before the weaker, nontested items, a category cued recall test is particularly suited to detecting interference effects.

Despite this, feedback did not affect retrieval-induced forgetting; instead, the data patterns

between Experiments 1 and 2 (which used recognition) were remarkably similar.

More broadly, our data contributes to the current debate about whether retrieval-induced

forgetting is better accounted for by interference or inhibition (e.g., Hulbert, Shivde, & Anderson,

2012; Raaijmakers & Jakab, 2013; Storm & Levy, 2012; Verde, 2012). According to both

inhibition and interference theorists, retrieval practice directly enhances representation of the

tested items. More controversial is what can cause the reduction in performance of the related,

nontested materials. According to interference accounts, retrieval-induced forgetting occurs

because the tested items block retrieval of the nontested items. Therefore, the very process that

strengthens the tested items causes retrieval deficits of the nontested items. An extension of this

reasoning is that variables that affect performance of the former should also affect performance

of the latter. In contrast, inhibition theorists argue that the testing effect and retrieval-induced

forgetting should be dissociable because they have different underlying mechanisms (with the

latter involving inhibition).

Of most relevance to the present study is the strength-independence assumption of the

inhibition theory, which suggests that magnitude of retrieval inhibition is independent of

memory strength of the tested items. This is consistent with the broader idea that different

mechanisms underlie the testing effect and retrieval-induced forgetting. For example, the

detrimental effects of retrieval-induced forgetting and the benefits of retrieval practice are

dissociable neurologically, such that they are accompanied by different electrophysiological

(Spitzer, Hanslmayr, Opitz, Mecklinger, & Bäuml, 2009) and hemodynamic responses (Wimber

et al., 2008). Moreover, several studies have shown that retrieval-induced forgetting does not

always happen with selective retrieval practice (Bauml & Kuhbandner, 2007; Chan, 2009; Román, Soriano, Gomez-Ariza, & Bajo, 2009).

Pertaining specifically to the strength-independence assumption, increasing encoding time for some studied items enhances recall of these items without affecting recall of the remaining items (e.g., Bauml, 1997). More recently, researchers have shown that retrieval-induced forgetting can occur even without selective strengthening of the tested items. For example, Storm and colleagues showed similar levels of retrieval-induced forgetting regardless of whether one can successfully recall the tested items during retrieval practice (Storm et al., 2006; by giving subjects studied/possible or nonstudied/impossible word stem, Storm & Nestojko, 2010). Similarly, Bauml (2002) showed that generating nonstudied items from the studied categories can cause retrieval-induced forgetting (see also Storm, Bjork, & Bjork, 2008). Both of these procedures, however, suffer from the same potential concern. Namely, participants may covertly retrieve some of the studied items even if they are not required to. For example, when participants are presented with an impossible, and thus very difficult, retrieval cue (e.g., Weapon – Wo____ - successful retrieval is impossible because no weapon begins with "Wo"), they might attempt to recall the studied items and then realize that none of them fit the word stem. Likewise, when participants must generate nonstudied category exemplars, the studied exemplars may come to mind (because they are more accessible than the nonstudied exemplars) and are then withheld from output. As a result, support for strength-independence has been criticized as less clear cut than previously assumed (Raaijmakers & Jakab, 2013).

In the present study, we showed that a manipulation that greatly enhances performance of the tested items had virtually no impact on performance of the nontested items. In fact, even participants who benefitted the most from feedback showed the same amount of (or less)

retrieval-induced forgetting (.12 in Experiment 1 and .06 in Experiment 2) than participants who received no feedback at all (.15 in Experiment 1 and .07 in Experiment 2). Moreover, these results were obtained regardless of whether the retrieval environment was more (Experiment 1) and less (Experiment 2) susceptible to associative interference. These results are clearly consistent with the strength-independence assumption and pose a serious challenge to interference accounts of retrieval-induced forgetting.

Another interesting finding is that although immediate feedback greatly increases the probability of retrieval success during retrieval practice, it produced no appreciable difference to the magnitude of retrieval-induced forgetting. As a result, we found no evidence for the hypothesis that increasing retrieval fluency during retrieval practice would reduce the need for inhibition (see the motivation behind the immediate and delayed feedback manipulation in the Introduction). This finding is consistent with the idea that retrieval inhibition is determined ***not by*** accessibility of the tested items (an instantiation of strength-independence), but by competition arising from the nontested items (Anderson et al., 2000; an instantiation of competition-dependence, Gómez-Ariza, Fernandez, & Bajo, 2012).

Applied Implications

In this experiment, we purposefully equated the total time of the retrieval practice phase for all participants, so feedback represents a cost free method for enhancing retention – it increased retention of the tested materials without any other performance costs. In real-life situations, it would be much more difficult to control for the total time spent on learning. For example, providing feedback in class (e.g., when instructors return exams to the students) may require instructors to sacrifice some lecture time. But given that errors are almost never corrected spontaneously in the absence of feedback (Metcalfe & Kornell, 2007), and retrieval-

induced forgetting can occur without overt retrieval (e.g., a student receives an in-class quiz but does not produce a response, Storm & Nestojko, 2010), we believe that forfeiting some lecture time is a reasonable sacrifice.

In two experiments, we found that feedback timing (i.e., immediate or delayed) did not influence the magnitude of its benefit, whereas former research has shown that delaying feedback can sometimes increase its benefit (Butler et al., 2007). This discrepancy may reflect two important differences between the way feedback was administered in the present study and in Butler and colleagues' study. In the present study, immediate feedback was administered three times (as necessitated by the three retrieval practice trials), but delayed feedback was administered only once. Therefore, immediate feedback might have benefited from distributed, or spaced, learning (Delaney, Verkoeijen, & Spirgel, 2010). Another possible reason that we did not observe a delayed feedback benefit is that our retention interval was too short (20 min). Indeed, Butler et al. found that the benefits of delayed feedback over immediate feedback emerged after 1 week but not after 1 day.

Aside from a direct increase in recall of tested information, feedback can provide other benefits. Recent work suggests that previously inhibited memories might be relearned faster than the noninhibited ones (Storm et al., 2008). This finding, again, demonstrates the critical importance of feedback for learning.

Future Directions and Conclusions

This research represents a first attempt to address the influence of feedback on memory for the nontested materials; as such, several questions remain unanswered. For example, would our results generalize to other materials and retention intervals? Previous research suggests that coherent materials (like prose instead of word lists) and a long delay can sometimes offset the

detrimental effects of retrieval-induced forgetting (Chan, 2009; Chan, McDermott, & Roediger, 2006), so it would be interesting to examine the effects of feedback on the nontested materials under these conditions.

In conclusion, we have demonstrated that providing feedback can enhance memory performance of the tested material without any detectable cost on the nontested material. Given this knowledge, it appears that educators can feel even more confident in using testing and feedback to enhance learning.

References

Anderson, M. C. (2003). Rethinking interference theory: Executive control and the mechanisms

    of forgetting. *Journal of Memory and Language*, *49*, 415–445.

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.

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.

Aslan, A., & Bäuml, K.-H. T. (2011). Individual differences in working memory capacity predict

    retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, *37*(1), 264–269. doi:10.1037/a0021324

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories A

    replication and extension of the Connecticut category norms. *Journal of Experimental*

    Psychology, *80*(3, Pt.2), 1–46. doi:10.1037/h0027577

Bauml, K.-H. (1997). The list-strength effect: Strength-dependent competition or suppression?

    Psychonomic Bulletin & Review, *4*(2), 260–264.

Bauml, K.-H. (2002). Semantic generation can cause episodic forgetting. *Psychological Science*,

    13(4), 356.

Bauml, 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

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated

    studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*,

1118–1133.

Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, *13*(4), 273–281. doi:10.1037/1076-898X.13.4.273

Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 918–928.

Camp, G., Pecher, D., Schmidt, H. G., & Zeelenberg, R. (2009). Are independent probes truly independent? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 934–942.

Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language*, *61*(2), 153–170. doi:10.1016/j.jml.2009.04.004

Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, *135*(4), 553–571. doi:10.1037/0096-3445.135.4.553

Ciranni, M. A., & Shimamura, A. P. (1999). Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1403–1414.

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation*, *53*, 63–147.

Gómez-Ariza, C. J., Fernandez, A., & Bajo, M. T. (2012). Incidental retrieval-induced forgetting of location information. *Psychonomic Bulletin & Review*, *19*(3), 483–489.

doi:10.3758/s13423-012-0228-8

Hulbert, J. C., Shivde, G., & Anderson, M. C. (2012). Evidence Against Associative Blocking as

a Cause of Cue-Independent Retrieval-Induced Forgetting. *Experimental Psychology*, *59*(1),

11–21. doi:10.1027/1618-3169/a000120

Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective

feedback modulate the effect of testing on memory retention. *European Journal of Cognitive

Psychology*, *19*, 528–558.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*.

Providence, RI: Brown University Press.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.).

Mahwah, NJ: Lawrence Erlbaum.

McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). The Read-Recite-Review Study

Strategy Effective and Portable. *Psychological Science*, *20*(4), 516–522.

Metcalfe, J., & Kornell, N. (2007). Principles of cognitive science in education: The effects of

generation, errors, and feedback. *Psychonomic Bulletin & Review*, *14*(2), 225–229.

Raaijmakers, J. G. W., & Jakab, E. (2012). Retrieval-induced forgetting without competition:

Testing the retrieval specificity assumption of the inhibition theory. *Memory & Cognition*,

40(1), 19–27.

Raaijmakers, J. G., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status

of the inhibition theory for forgetting. *Journal of Memory and Language*, *68*, 98–122.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *16*(2), 163–178.

doi:10.1037/0278-7393.16.2.163

Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and

implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210.

Román, P., Soriano, M. F., Gomez-Ariza, C. J., & Bajo, M. T. (2009). Retrieval-induced

forgetting and executive control. *Psychological Science*, *20*(9), 1053–1058.

Spitzer, B., Hanslmayr, S., Opitz, B., Mecklinger, A., & Bäuml, K.-H. (2009). Oscillatory

Correlates of Retrieval-induced Forgetting in Recognition Memory. *Journal of Cognitive

Neuroscience*, *21*(5), 976–990. doi:10.1016/j.neuroimage.2003.09.034

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

Storm, B. C., & Nestojko, J. F. (2010). Successful inhibition, unsuccessful retrieval:

Manipulating time and success during retrieval practice. *Memory*, *18*(2), 99–114.

doi:10.1080/09658210903107853

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated relearning after retrieval-induced

forgetting: The benefit of being forgotten. *Journal of Experimental Psychology: Learning,

Memory, and Cognition*, *34*(1), 230–236. doi:10.1037/0278-7393.34.1.230

Storm, B. C., Bjork, E. L., Bjork, R. A., & Nestojko, J. F. (2006). Is retrieval success a necessary

condition for retrieval-induced forgetting? *Psychonomic Bulletin & Review*, *13*(6), 1023–

1027.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York:

Teachers College, Columbia University.

Tulving, E., & Hastie, R. (1972). Inhibition effects of intralist repetition in free recall. *Journal of

Experimental Psychology: Learning, Memory, and Cognition*, *92*(3), 297–304.

doi:10.1037/h0032367

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the

    operation span task. *Behavioral Research Methods*, *37*, 498–505.

Veling, H., & van Knippenberg, A. (2004). Remembering Can Cause Inhibition: Retrieval-

    Induced Inhibition as Cue Independent Process. *Journal of Experimental Psychology:*

    Learning, Memory, and Cognition, *30*(2), 315–318. doi:10.1037/0278-7393.30.2.315

Verde, M. F. (2009). The list-strength effect in recall: Relative-strength competition and retrieval

    inhibition may both contribute to forgetting. *Journal of Experimental Psychology: Learning,*

    Memory, and Cognition, *35*(1), 205–220. doi:10.1037/a0014275

Verde, M. F. (2012). Retrieval-induced forgetting and inhibition: A critical review. In B. H. Ross

    (Ed.), *Psychology of Learning and Motivation* (Vol. 56, pp. 47–80). New York: Academic

    Press. doi:10.1016/B978-0-12-394393-4.00002-9

Williams, C. C., & Zacks, R. T. (2001). Is retrieval-induced forgetting an inhibitory process?

    American Journal of Psychology, *114*(3), 329–354.

Wimber, M., Bäuml, K.-H., Bergström, Z., Markopoulos, G., Heinze, H.-J., & Richardson-

    Klavehn, A. (2008). Neural markers of inhibition in human memory retrieval. *Journal of*

    Neuroscience, *28*(50), 13419–13427. doi:10.1523/JNEUROSCI.1916-08.2008

Yonelinas, A. P., Hockley, W. E., & Murdock, B. B. (1992). Tests of the list-strength effect in

    recognition memory. *Journal of Experimental Psychology: Learning, Memory, and*

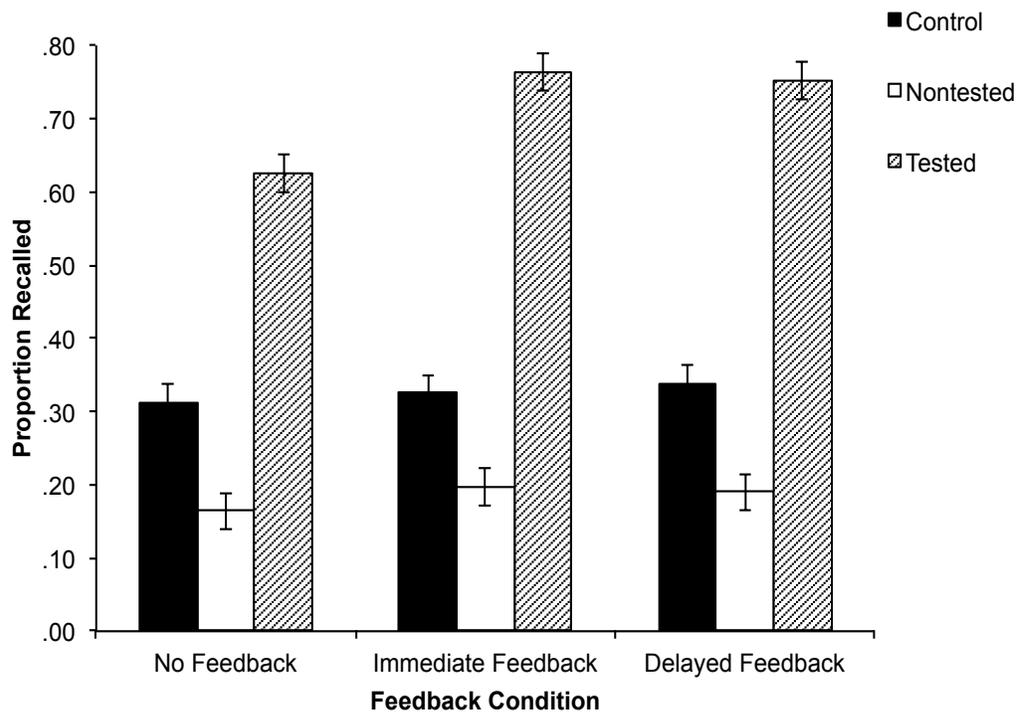   *Cognition*, *18*(2), 345–355. doi:10.1037/0278-7393.18.2.345

Figure 1. Proportion of exemplars correctly recalled on the category cued recall final test in

Experiment 1. Error bars reflect the within-subjects 95% confidence interval.
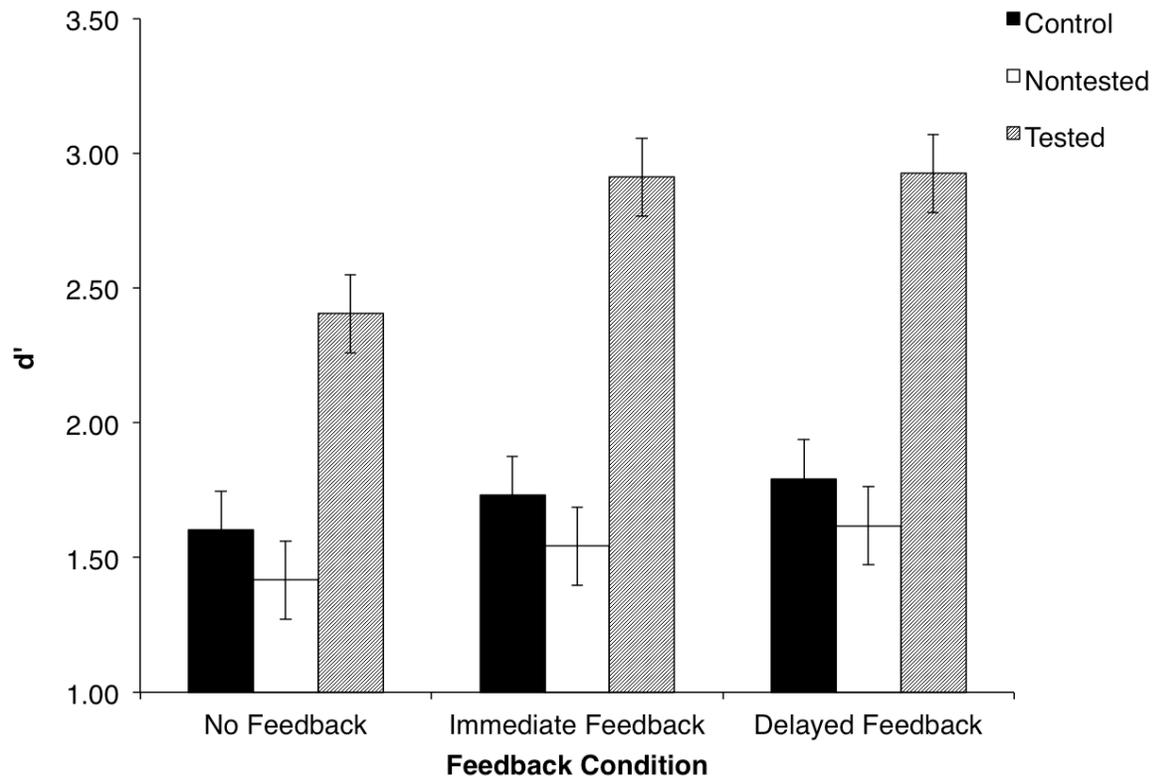
Figure 2. Recognition performance in Experiment 2 as displayed by *d'*. Error bars reflect the within-subjects 95% confidence interval.

Appendix

Materials used in the experiments. Asterisks denote exemplars used in Experiment 1. Note that

the category "Religion" in Experiment 1 was replaced by the category "Metal" in Experiment 2.

Category Name		Exemplar Word	Category Typicality	Kucera-Francis Word Frequency	Thorndike-Lorge Word Frequency
Birds	*	Canary	.30	N/A	39
	*	Blackbird	.20	N/A	4
	*	Oriole	.17	5	5
	*	Parrot	.16	1	18
	*	Starling	.11	N/A	7
	*	Vulture	.10	1	12
		Pigeon	.13	3	76
		Woodpecker	.10	1	6
		Swallow	.09	10	170
		Falcon	.06	4	5
		Flamingo	.04	N/A	2
		Raven	.03	N/A	9
Job	*	Dentist	.25	12	61
	*	Plumber	.09	4	28
	*	Accountant	.08	2	18
	*	Chemist	.05	16	32
	*	Janitor	.03	4	15
	*	Baker	.03	36	86
		Engineer	.25	42	218

		Clerk	.07	34	226
		Policeman	.05	19	190
		Merchant	.05	20	118
		Fireman	.05	1	45
		Architect	.04	22	102
Animals	*	Mouse	.27	10	48
	*	Giraffe	.19	N/A	1
	*	Squirrel	.13	1	32
	*	Donkey	.11	1	20
	*	Rabbit	.11	11	96
	*	Leopard	.10	N/A	13
		Tiger	.46	7	103
		Elephant	.41	7	144
		Sheep	.19	23	86
		Camel	.06	1	24
		Antelope	.06	7	49
		Turtle	.03	8	21
Flowers	*	Pansy	.24	6	39
	*	Gardenia	.20	1	20
	*	Daffodil	.20	N/A	17
	*	Lilac	.10	4	52
	*	Sunflower	.08	N/A	14
	*	Marigold	.04	N/A	10

		Violet	.33	7	109
		Orchid	.31	N/A	44
		Peony	.09	3	N/A
		Azalea	.07	2	N/A
		Poppy	.05	2	16
		Buttercup	.04	N/A	8
Weather	*	Lightning	.16	14	101
	*	Blizzard	.10	7	27
	*	Monsoon	.10	3	2
	*	Sunny	.07	13	53
	*	Humidity	.07	8	17
	*	Drought	.04	5	23
		Typhoon	.13	1	N/A
		Thunder	.12	14	138
		Shower	.07	15	147
		Hailstorm	.04	1	N/A
		Rainbow	.03	4	45
		Snowstorm	.03	3	11
Spice	*	Cinnamon	.27	N/A	82
	*	Paprika	.20	2	72
	*	Mustard	.14	20	51
	*	Nutmeg	.10	4	21
	*	Thyme	.07	N/A	7

	*	Ginger	.03	2	68
		Vanilla	.27	1	86
		Garlic	.27	4	30
		Cloves	.21	1	N/A
		Onions	.14	15	N/A
		Vinegar	.09	9	92
		Chives	.03	1	N/A
Cloth	*	Madras	.07	N/A	7
	*	Burlap	.07	N/A	18
	*	Denim	.07	N/A	17
	*	Flannel	.06	4	79
	*	Tweed	.03	5	76
	*	Canvas	.03	19	61
		Satin	.18	5	116
		Velvet	.07	4	177
		Corduroy	.06	1	13
		Acetate	.04	7	N/A
		Jersey	.03	25	106
		Chiffon	.02	N/A	78
Religion	*	Altar	.01	5	58
	*	Parish	.03	11	41
	*	Shrine	.10	7	33
	*	Cathedral	.38	8	56

*	Tabernacle	.05	1	5
*	Mosque	.26	10	9
Metal	Bronze	.18	11	65
	Platinum	.17	4	35
	Nickel	.15	7	39
	Mercury	.07	10	39
	Uranium	.05	6	N/A
	Tungsten	.05	4	N/A
	Cobalt	.03	2	5
	Titanium	.04	2	N/A
	Chromium	.03	4	10
	Cadmium	.02	2	N/A
	Magnesium	.14	N/A	8
	Sodium	.09	12	4