Which Route to Recovery? Controlled Retrieval and Accessibility Bias in Retroactive Interference
Author(s): Cindy Lustig, Alex Konkel and Larry L. Jacoby
Source: Psychological Science, Vol. 15, No. 11 (Nov., 2004), pp. 729-735
Published by: Sage Publications, Inc. on behalf of the Association for Psychological Science
Stable URL: http://www.jstor.org/stable/40064038
Accessed: 25-10-2017 18:54 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://about.jstor.org/terms
Research Article

Which Route to Recovery?
Controlled Retrieval and Accessibility Bias in Retroactive Interference

Cindy Lustig, Alex Konkel, and Larry L. Jacoby

Washington University

ABSTRACT—New learning often interferes with the production of older, previously learned responses. However, the original responses usually appear to spontaneously recover and regain their dominance after a delay. This article takes a new approach to questions of interference and recovery by examining performance on immediate and delayed tests using direct or indirect instructions. Direct instructions asked participants to deliberately retrieve the original responses, and indirect instructions allowed them to respond on a more automatic basis, using whatever response came to mind first. Results suggest that interference and recovery may have their largest effects via relatively automatic influences on memory, such as the accessibility of new versus original information. This finding adds a new perspective to classic theories of interference and recovery, and may also inform current understanding of performance in populations (e.g., older adults) that often rely predominantly on automatic memory processing.

People usually think of memories as fading as time goes by. In addition, learning new information often interferes with the retrieval of older memories. At the same time, old habits are infamous for their ability to return. Both the retroactive interference (RI) caused by new learning and the spontaneous recovery of old information after a delay have been observed at least since the classic experiments of Pavlov (1927), but how they occur remains a mystery. In this report, we take a new perspective on the problem by asking how time affects both the controlled, deliberate retrieval of old memories and their accessibility.

RI and spontaneous recovery are typically studied by first teaching participants one response to a cue (e.g., a tone signals the delivery of food; say bend when presented with knee), and then teaching them a new response (e.g., the tone now signals the delivery of a shock; say bone in response to knee). RI is shown when a group that learns both responses shows poorer memory for the originally learned response compared with a control group that learned only the original response. RI is usually strongest on an immediate test; over time, the RI group’s production of the original responses becomes more similar to that of the control group, a phenomenon known as spontaneous recovery. Today, RI and recovery are studied not only in the context of animal learning, but in diverse areas of inquiry, including misinformation effects in courtroom testimony (Loftus, 1975; McCloskey & Zaragoza, 1985), and as potential bases for recidivism following treatment for phobias, addictions, and other disorders (e.g., Bouton, 2000, 2002).

Classic interference theory in human learning and memory centered on two factors, unlearning and response competition (see reviews by Bjork, 2001; Brown, 1976; Crowder, 1976; Postman, 1971; Postman & Underwood, 1973; Wheeler, 1995). Unlearning hypotheses suggested that learning the new response directly weakens the association between the cue and the original response, and that over time the original association recovers from this blow (Melton & Irwin, 1940; Underwood, 1971). Response-competition hypotheses suggested that RI occurs because the new response competes with, and perhaps temporarily suppresses the retrieval of, the original response (McGeogh, 1942; Postman, Stark, & Fraser, 1968). This suppression was assumed to be temporary, and its dissipation reallowed the retrieval of old information. Although interference theory was extremely influential, being described in Crowder’s (1976) classic memory textbook as one of the “most extensive theoretical efforts within all of experimental psychology” (p. 217), attempts to explain all the relevant data on the basis...
of unlearning and response competition became increasingly complicated. Ultimately, many investigators in human memory abandoned interference theory to pursue other topics during the “cognitive revolution.”

Here we take a new approach to questions of RI and recovery, one that complements classic two-factor interference theory, by taking advantage of one of the cognitive revolution’s main insights: Memory is not a unitary construct, but rather the complex outcome of multiple systems or processes, some controlled and some automatic, that may make independent contributions (Jacoby, 1991, 1998; Schacter, 1987). For example, controlled processes include an active search through memory for the original response, or a conscious effort to constrain responses to those that were learned first. In contrast, automatic processes could include a simple bias to say the new, second-learned response, because it has occurred most recently and is the first to come to mind.

Whereas classic two-factor theory focused on unlearning and response competition as mechanisms of interference and recovery, this two-process approach contrasts the roles of controlled and automatic processing. It suggests that RI, or at least the tendency to produce the new response, may often have its effects via automatic processes, such as guessing (Jacoby, Hessels, & Bopp, 2001). That is, if the participant does not retrieve the original response through a controlled, recollective search, he or she may instead produce the recently learned new response because it is still “fresh in mind” and highly accessible. A series of experiments using the process-dissociation procedure (Jacoby, 1991, 1998) found that exposure to interfering information (e.g., seeing knee-bone after learning knee-bend) influenced estimates of accessibility bias only, leaving estimates of controlled processing unaffected (Jacoby et al., 2001). Studies using tests designed to minimize controlled retrieval often produce results consistent with this finding, showing large effects of interference (see review by Lustig & Hasher, 2001a). Furthermore, amnesic and frontal lobe patients, whose controlled retrieval abilities are impaired because of brain injury, have an increased vulnerability to interference (e.g., Shimamura, Jurica, Mangels, Gershberg, & Knight, 1995; Winocur, Moscovitch, & Bruni, 1996). Taken together, these results suggest that interfering information often has its strongest effects via automatic influences on memory.

What about spontaneous recovery? To what degree is it due to changes in controlled, recollective retrieval versus changes in more automatic influences? We addressed this question in the present study by using the materials from Jacoby et al. (2001) in a paired-associate list RI procedure (i.e., learn knee-bend, then learn knee-bone) and then giving participants immediate and delayed tests using either direct or indirect instructions.

Participants given direct test instructions were explicitly told to constrain their responses to the original list, and that responses from the new, second list would not be considered correct answers. Therefore, in the direct test, participants were expected to attempt controlled, recollective retrieval of the original responses. Any production of the new responses would reflect failures of this controlled retrieval attempt and a resort to more automatic processes, such as guessing.

With the direct test, we expected to replicate standard findings of relative recovery, the form of spontaneous recovery found most frequently in the literature. (See discussions by Brown, 1976, and Postman et al., 1968.) That is, although the rate of original responses may not significantly increase over the delay in the RI condition (in absolute terms), there should be less forgetting of the original responses in the RI condition than in the control condition. In our view, this pattern occurs because information supporting controlled retrieval of the original responses deteriorates over the delay (affecting performance in both the control and RI conditions), whereas accessibility influences on guessing change to favor the original responses over the new, alternate responses (affecting only the RI condition).

The indirect test was included to provide insight into changes in accessibility. For this test, participants were simply told to produce the first response that came to mind, and thus had no reason to attempt controlled, deliberate retrieval of one response in favor of the other. For the indirect test, we were most interested in the relative rate of production for the original responses versus the new, alternate responses. We expected that the alternate responses would be quite accessible immediately after they were learned, and thus the rate of producing alternate responses would be relatively high on a test taken at that time. However, such a recency advantage should fade over time (cf. Estes, 1955), so we expected that on the delayed test, the original responses would become dominant and, therefore, be produced at a higher rate than the alternate responses. As we have already described, changes in accessibility likely also influence performance on direct tests: When deliberate retrieval fails, participants may rely on guessing, with their guessed responses determined by which response is more accessible. Changes in the automatic accessibility of original versus alternate responses may thus play an important role in spontaneous recovery.

**METHOD**

**Participants, Design, and Materials**

Participants were 48 Washington University undergraduates, randomly assigned to the direct or indirect testing conditions. Data from 1 additional participant were dropped because of experimenter error. Participants received either course credit or $15 as compensation.

Instructions (direct, indirect) were manipulated between subjects. Interference (control, RI), test time (immediate, delayed), and response type (original, alternate) were within-subjects variables. The dependent variable was the rate at which each response type (original, alternate) was produced.

The materials and basic design of the study are outlined in Table 1. Materials were modified from lists used by Hay and Jacoby (1996, 1999). These materials reliably produce

---

**Accessibility and Interference**

---
interference effects (see also Jacoby, 1994; Jacoby et al., 2001), although they differ somewhat from those used in many verbal learning experiments in that they use related, rather than unrelated, cue-response pairs. Each participant studied two lists that included items from both the control and RI conditions; half the items in each condition were designated to be tested immediately after the study phase, and the other half were tested the next day.

Each study list began and ended with three buffer word pairs that were not tested. Between the buffer pairs were 40 critical pairs, with items from the two interference conditions and two test-time conditions randomly intermixed, with the restriction that no more than three pairs of one type occurred consecutively. Each word pair consisted of a cue word (e.g., knee) plus a semantically related response word. The response words for the first study list were the target words for the later fragment tests (e.g., bend was paired with knee and was the target for the later test item knee-b_n_). In the RI condition, the response words for the second study list were orthographically similar to the response words for the first study list (e.g., bone was paired with knee), and were therefore alternate completions for the fragment tests. In the control condition, the response words for the second study list were not orthographically similar to the response words in the first study list and could not complete the fragments in the tests (e.g., jerk was paired with knee).

Each test pair consisted of a cue word and a word fragment. Cue-fragment pairs were chosen such that the two orthographically similar response words (e.g., bend and bone) would be elicited with equal probability, according to previous norms. Each test list began and ended with three buffer pairs that did not correspond to any of the pairs seen at study. In between the buffer pairs were 30 critical pairs: 10 pairs corresponding to study-list pairs in the control condition (e.g., knee-bend/jerk), 10 pairs corresponding to study-list pairs in the RI condition (e.g., knee-bend/bone), and 10 pairs that were not on the study list (baseline pairs). These baseline pairs were included to provide a measure of chance performance on the fragment-completion test. Control, RI, and baseline pairs were randomly intermixed with the restriction that no more than 3 pairs of any one type occurred consecutively. Materials were counterbalanced across participants so that the two possible completions (e.g., bend and bone) of each fragment served equally often as targets across the control, RI, and baseline test conditions and the two test days.

Participants were tested individually on a computer using E-Prime software. Word and word-fragment pairs were presented in the middle of the screen in lowercase letters. Each study-list trial began with a 500-ms fixation cross, followed by the word pair presented for 3 s. Each fragment-test trial began with a 500-ms fixation cross, followed by the word-fragment pair presented for up to 5 s.

### Procedure

After completing informed-consent procedures and a demographics questionnaire, participants were told that they would be viewing pairs of related words (e.g., coffee-cup) for 3 s per pair. They were asked to remember each pair for a later memory test, and instructed to use the presentation time to think about the association between the two words. They next completed three practice trials, were given the opportunity to ask questions, and began the first study list.

After the first study list, participants completed the Trail-Making Test (Reitan, 1958) as a nonverbal filler task and were then given instructions for the second study list. They were told that they would once again be presented with pairs of related words for 3 s and asked to think about the relation between them, but that although the first word in each pair would be the same as in the original study list, the second word would differ. For example, if they had learned the pair coffee-cup in the original list, now they might see coffee-mug. Participants were also told that it was very important to learn this second list carefully, because on the later memory test correct answers might come either from this list or from the original list. They were given the opportunity to ask questions and then began.

Immediately after completing the second study list, participants were given instructions for the fragment test. They were shown two examples of word-fragment pairs (e.g., business-de__) and told that their task was to complete each fragment with a word related to the first word. Participants in the direct-testing condition were told that if the cue word was a new word that they had not seen at study, they should complete the fragment with the first word that came to mind. However, if the cue word was

### TABLE 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>List 1</th>
<th>List 2</th>
<th>Day 1 test</th>
<th>Day 2 test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 1 test, control</td>
<td>bed-sheet</td>
<td>bed-linen</td>
<td>bed-s_e_e_</td>
<td>—</td>
</tr>
<tr>
<td>Day 1 test, retroactive interference</td>
<td>knee-bend</td>
<td>knee-bone</td>
<td>knee-b_n_</td>
<td>—</td>
</tr>
<tr>
<td>Day 1 test, baseline</td>
<td>—</td>
<td>—</td>
<td>limp-w__k</td>
<td>—</td>
</tr>
<tr>
<td>Day 2 test, control</td>
<td>carry-load</td>
<td>carry-drop</td>
<td>—</td>
<td>carry-_o_d</td>
</tr>
<tr>
<td>Day 2 test, retroactive interference</td>
<td>rent-house</td>
<td>rent-lease</td>
<td>—</td>
<td>rent-_ _ _se</td>
</tr>
<tr>
<td>Day 2 test, baseline</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>belly-f_o_</td>
</tr>
</tbody>
</table>
one they had seen in the study lists, they should say the response word paired with it on the original study list; words from the second list would not be counted as correct answers. Participants in the indirect-testing condition were told to complete each fragment with the first related word that came to mind. All participants were told that they had up to 5 s to complete each trial, but could terminate the trial earlier by saying their response and pressing the space bar. The fragment test completed testing for the first day.

Participants returned the next day for the second fragment test and then completed the Shipley (1940) and Extended Range (Educational Testing Service, 1976) vocabulary tests.

RESULTS

The scores of the two instruction groups did not differ on either the Shipley vocabulary test (direct: $M = 32.0$, $SE = 0.5$; indirect: $M = 32.2$, $SE = 0.8$) or the Extended Range Vocabulary Test (direct: $M = 22.3$, $SE = 1.6$; indirect: $M = 20.7$, $SE = 1.5$), both $t$ < 1. Completion rates for the baseline cue-fragment pairs, for which the participants had not studied a solution, were equivalent ($M_s = .35$) for the two orthographically similar response words for each item (e.g., bend and bone). An Instructions $\times$ Test Time $\times$ Response Type analysis of variance (ANOVA) on frequency of baseline responses did not reveal any main effects or interactions, all $ps > .10$. Thus, chance performance was equivalent across conditions.

We were interested in the production of original and alternate responses across instructions (direct, indirect), interference conditions (control, interference) and test time (immediate, delayed). Means and standard errors are displayed in Figure 1. Given the large number of comparisons possible in this design, we restricted our analyses to those directed to our specific questions.

As a first step, it was important to establish that our design and procedures replicated standard findings of RI and spontaneous recovery on direct memory tests. The originally learned responses from participants in the direct-instructions condition were thus submitted to an Interference $\times$ Test Time ANOVA, yielding a significant interaction, $F(1, 23) = 7.31, p < .05, f^2 = 0.30$. Immediate testing resulted in interference on the direct test, with more frequent production of the original responses in the control than in the RI condition, $t(23) = 3.08, p < .01, d = 0.63$. However, after the delay, performance was equivalent in these two conditions, $t < 1$. The indirect test showed a strong effect of interference, $F(1, 23) = 14.69, p < .001, f^2 = 0.62$, that did not change over the delay, $F < 1$, replicating other findings of interference (e.g., Lustig & Hasher, 2001b; Nelson, Keelean, & Negrao, 1989) and delay-maintained performance (e.g., Sloman, Hayman, Ohta, Law, & Tulving, 1988, Experiment 2) on indirect tests.

Notably, the largest change in frequency of original responses was their decrease over the delay in the control condition of the direct test. In contrast, the production of original responses in the RI condition was relatively stable for this test. These results are thus an example of relative recovery: Although the rate of original responses within the RI condition does not significantly increase over the delay, it shows less forgetting than would be predicted on the basis of the control condition (see discussions by Brown, 1976; Postman et al., 1968). The most likely explanation for the control condition’s decrease in original responses over time on the direct test is a delay-related decline in information supporting controlled, deliberate retrieval. The fact that performance in the control conditions of the direct and indirect tests was nearly identical after the delay is consistent with this explanation.

For the indirect test, we were most interested in how the production of original responses versus newly learned, alternate responses changed over the delay in the RI condition. That is, given that the participant had learned two potential responses, how did the relative dominance of those responses change over time? The idea that learning a new response has its primary effect by providing an easily accessible alternative when the original response is not retrieved (either via controlled recollection or, in this case, by being the first to come to mind)

---

Original and alternate responses did not completely constrain each other. Participants could fail to respond or give an extra-experimental response (e.g., bang rather than bend or bone). These types of responses occurred at rates of 8 to 21%, respectively, across conditions ($M = 14\%$). Thus, the significant interaction is not obligated by our design; one might have predicted only a main effect of delay, with both original and alternate response types returning to baseline.
predicts that the rate of producing new responses should be especially sensitive to the delay: Their accessibility advantage, and thus their production, should decrease over the delay (Estes, 1955; see Bjork, 2001, for a recent discussion of the generality of this phenomenon).

This was indeed the case. Within the RI condition for the indirect test, there was a significant interaction between test time and response type, $F(1, 23) = 6.69, p < .05, f^2 = 0.28$. Secondary analyses confirmed the visual impression (see Fig. 1) that although the frequency of original responses changed very little or even increased over the delay, $t(23) = 1.64, p = .11$, the frequency of alternate responses decreased, $t(23) = 3.19, p < .005, d = 0.65$. Although the effect was numerically larger for the indirect test condition, an analysis including the direct test also found a significant interaction between test time and response type, $F(1, 46) = 7.76, p < .01, f^2 = 0.16$, that did not interact further with instructions, $F < 1$. Thus, changes in the rates of original versus alternate responses followed similar patterns for the direct and indirect tests, suggesting that both were influenced by changes in accessibility bias.

**DISCUSSION**

How does new learning interfere with the production of earlier-learned responses, and why do those original responses seem to recover over time? In the present study, we took a new look at questions of RI and recovery by asking how they might result from the interplay of controlled and automatic influences on memory. Our results suggest that a major route for RI and recovery is automatic, reflecting changes in accessibility.

Our concept of how the automatic accessibility of original and new responses may change over time has much in common with an explanation for RI and recovery first described by Miller and Stevenson (1936), and later discussed in Crowder’s (1976) classic textbook on memory (see also Brown, 1976). Figure 2 reproduces Crowder’s illustration. As depicted in the figure, forgetting is rapid but negatively accelerated, and thus new learning may occur when the strength (in our terms, accessibility) of original learning has already reached a relatively flat portion of its forgetting curve. On a test that immediately follows new learning, the strength of the new response will be relatively high, and thus this response will be produced at a rate equal to or greater than the rate for the original response. As the strength (accessibility) of the new response begins to decline, it will eventually become lower than that of the original response.

Why do original responses maintain their strength better than new responses over long delays? The reasons are still uncertain, although the phenomenon has been one of the fundamentals of memory research since Jost’s (1897) law. One proposal is that participants use their “best,” or most elaborative, encodings when learning the original list, and do not encode later lists as well (Hasher, Griffin, & Johnson, 1977; Hasher & Johnson, 1975; see Bouton, 2000, 2002, for a related explanation applied to animal learning). On the immediate test, production of the new responses is high because they have just been learned and are thus very accessible, or “fresh in mind.” However, this advantage fades quickly (e.g., Estes, 1955), eventually allowing the greater encoding strength of the original responses to be revealed.

Previous discussions of changes in strength and their influence on interference phenomena described memory as a unitary construct, without reference to the distinction between controlled and automatic processes. However, many classic interference experiments used instructions similar to those of our indirect test, asking participants for the first response that came to mind and thus minimizing deliberate, targeted retrieval from one particular list (e.g., Underwood, 1948). This practice was adopted to build on studies of conditioning extinction and recovery in animals (e.g., Miller & Stevenson, 1936), who are of course unlikely to engage in deliberate retrieval. Influences on automatic processes may thus be responsible for interference and recovery effects across many situations.

Most present-day memory tests with humans use instructions similar to those of our direct test; that is, they instruct participants to engage in controlled, recollective retrieval of a particular item. Although direct tests include a controlled retrieval component, changes in automatic influences such as accessibility also affect direct tests, in cases in which controlled retrieval is attempted, but fails (Jacoby, 1991; Jacoby et al., 2001). If controlled retrieval fails, participants may respond on the basis of guessing or familiarity. Our results, in combination with those of previous experiments using the process-dissociation procedure (Hay & Jacoby, 1996, 1999; Jacoby, 1991; Jacoby et al., 2001), suggest that response competition may have its largest effects through these relatively automatic processes.
rather than by directly disrupting controlled, recollective retrieval processes.

Focusing on the different roles of automatic versus controlled processes may aid in understanding not only interference per se, but also the memory performance of populations that are unlikely or unable to engage the controlled processes that can help minimize interference. Beyond the obvious case of animal learning, automatic influences likely play a predominant role in the memory performance of patients whose brain damage impairs their ability to engage in controlled processing (e.g., Shimamura et al., 1995; Winocur et al., 1996). Likewise, older adults often have impaired controlled processing and rely largely on automatic processing (Hasher & Zacks, 1979; Jennings & Jacoby, 1993). In light of this reliance on automatic memory processes, might the present results help explain why it can be difficult to teach an old (or damaged) brain "new tricks"—or at least retain them in memory?

Many other questions about interference and recovery remain. For example, how might changes in accessibility contribute to performance on so-called modified-modified-free-recall tests that ask the participant for both responses (Barnes & Underwood, 1959), or how might such changes contribute to absolute recovery? Which interference variables act primarily through automatic processes, and which through controlled processes (cf. Yonelinas & Jacoby, 1994)? We suggest that the answers to these questions may be best found by considering that memory is not a unitary construct, but rather a complex phenomenon influenced by both controlled and automatic processes. Exploring the roles of recollective retrieval and automatic accessibility in interference may help explain why, if people are not vigilant to keep them under control, their old bad habits creep back over time.

Acknowledgments—We thank Martha Storandt for statistical advice. This research was supported by National Institute on Aging (NIA) Grant AG13845 to Larry Jacoby. Cindy Lustig was supported by NIA Training Grant AG00030 and the Alzheimer Association.

REFERENCES


(Received 9/3/03; Revision accepted 10/31/03)